

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2005, 22:38:28 ; Search time 1549 Seconds
(without alignments)
10618.584 Million cell updates/sec

Title: US-10-751-550-1
Perfect score: 2690
Sequence: 1 tctagataatagactact.....ccataactagatcacaccatgg 2690

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5654200 seqs, 3057283753 residues

Total number of hits satisfying chosen parameters: 11308400

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*

3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*

10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*

11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*

12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*

13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*

14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*

15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*

16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*

17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*

18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*

19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*

20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*

21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2690	100.0	2690	19	US-10-751-550-1
2	843.8	31.4	387	18	US-10-845-059-10
3	192.4	7.2	4830	18	US-10-437-963-1638
4	160.4	6.0	4920	18	US-10-425-115-54556
5	160.4	6.0	7593	18	US-10-425-115-159798
6	158.8	5.9	6624	18	US-10-425-115-54546
7	157.2	5.8	3975	18	US-10-425-115-54557
8	156.8	5.8	1031	18	US-10-437-963-4747
9	156.2	5.8	1126	18	US-10-437-963-85940
10	155.6	5.8	5913	18	US-10-425-115-159738
11	155	5.8	4014	18	US-10-437-963-25937

c	12	155	5.8	7495	18	US-10-437-963-26057	Sequence 26057, A
	13	154.6	5.7	5928	18	US-10-437-963-85935	Sequence 85935, A
	14	154	5.7	5460	18	US-10-425-115-54555	Sequence 54555, A
	15	154	5.7	7170	18	US-10-425-115-159747	Sequence 159747, A
	16	153.4	5.7	4884	18	US-10-437-963-17763	Sequence 17763, A
	17	152.4	5.7	3792	18	US-10-425-115-159762	Sequence 159762, A
	18	152.4	5.7	4017	18	US-10-425-115-159761	Sequence 159761, A
	19	151.8	5.6	6108	18	US-10-437-963-26000	Sequence 26000, A
	20	151.4	5.6	2000	17	US-10-260-238-2549	Sequence 2549, Ap
	21	150.8	5.6	5547	18	US-10-425-115-54553	Sequence 54553, A
	22	150.2	5.6	4398	18	US-10-437-963-25899	Sequence 25899, A
	23	150.2	5.6	5700	18	US-10-437-963-25816	Sequence 25816, A
	24	149.2	5.5	4797	17	US-10-260-238-1495	Sequence 1495, Ap
	25	149.2	5.5	4989	18	US-10-437-963-25811	Sequence 25811, A
	26	148.8	5.5	5548	18	US-10-437-963-25978	Sequence 25978, A
	27	148.6	5.5	4425	18	US-10-437-963-26103	Sequence 26103, A
	28	148.6	5.5	4434	18	US-10-437-963-26124	Sequence 26124, A
	29	147	5.5	4860	18	US-10-437-963-26015	Sequence 26015, A
c	32	146	5.4	4272	18	US-10-425-115-159736	Sequence 159736, A
	31	146	5.4	3144	18	US-10-425-115-159765	Sequence 159765, A
	32	146	5.4	5901	18	US-10-425-115-159736	Sequence 159736, A
	33	145.8	5.4	6866	18	US-10-425-115-54545	Sequence 54545, A
	34	145.4	5.4	3615	18	US-10-437-963-25886	Sequence 25886, A
	35	144.2	5.4	1089	18	US-10-425-115-166297	Sequence 166297, A
	36	144.2	5.4	918	18	US-10-437-963-25867	Sequence 25867, A
	37	143.8	5.3	4809	18	US-10-425-115-159746	Sequence 159746, A
	38	142.6	5.3	5529	18	US-10-437-963-26012	Sequence 26012, A
c	39	140.8	5.2	7567	18	US-10-437-963-26107	Sequence 26107, A
	40	139	5.2	4557	18	US-10-437-963-26006	Sequence 26006, A
	41	136.6	5.1	6209	18	US-10-425-115-159776	Sequence 159776, A
	42	136.6	5.1	4755	18	US-10-437-963-25893	Sequence 25893, A
	43	134.2	5.0	5523	18	US-10-437-963-25999	Sequence 25999, A
	44	133.6	5.0	5157	18	US-10-437-963-25829	Sequence 25829, A
	45	132.6	4.9				

ALIGNMENTS

RESULT 1

US-10-751-550-1
; Sequence 1, Application US/10751550
; Publication No. US20050034192A1
; GENERAL INFORMATION:
; APPLICANT: The Texas A&M University System
; TITLE OF INVENTION: STEM-REGULATED, PLANT DEFENSE PROMOTER
; FILE REFERENCE: 017575.0775
; CURRENT APPLICATION NUMBER: US/10751,550
; CURRENT FILING DATE: 2004-01-05
; PRIOR APPLICATION NUMBER: US 60/437,974
; PRIOR FILING DATE: 2003-01-03
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2690
; TYPE: DNA
; ORGANISM: Sugarcane
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)...(2686)
US-10-751-550-1

Query Match 100.0%; Score 2690; DB 19; Length 2690;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2690; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTAGATATACGACTCTACTATAGGCGACGCGTGGTCGACGCCCGGGCTGCTCGGAC 60

Db 1 TCTAGATATACGACTCTACTATAGGCGACGCGTGGTCGACGCCCGGGCTGCTCGGAC 60

Qy 61 ACCTAGAGCGCCACCGCGTCTCTAGCTTCTCCAACTTCTGTCGAGATCCCTTCAGGG 120

Db 61 AGCTAGAGCGCCACCGCTCCTAGCTTCTCCAACTTCTCGTCGAGATCCCTTCAGGG 120
QY 121 ATGCCCAATGCGACCGCCCTTAAGTCAACTCTGCGGAGCTGAGCTTTCGCGAGGGTCAGA 180
Db 121 ATGCCCAATGCGACCGCCCTTAAGTCAACTCTGCGGAGCTGAGCTTTCGCGAGGGTCAGA 180
QY 181 GCTGCGGAGCAGCCCTGCTGAGCCGATTCCTGATGACCGCGGGGTGCGCTCCATGAAG 240
Db 181 GCTGCGGAGCAGCCCTGCTGAGCCGATTCCTGATGACCGCGGGGTGCGCTCCATGAAG 240
QY 241 AAGTGCATTTCGCCCAACCAAGTCGAGTGGGTGCGCTGAGGGGGCGGGAAGCAAAACGT 300
Db 241 AAGTGCATTTCGCCCAACCAAGTCGAGTGGGTGCGCTGAGGGGGCGGGAAGCAAAACGT 300
QY 301 TGCATGCACCTAGCGCCCTGCGAGCGAGCTCTGTAGTATCACTCGCTCGCTCCAGCT 360
Db 301 TGCATGCACCTAGCGCCCTGCGAGCGAGCTCTGTAGTATCACTCGCTCGCTCCAGCT 360
QY 361 CATGCTCGCAAGCTCTCAGGGGGCGCGCAGTGTCTCCAACTTTTCGCGCTCCTCTACA 420
Db 361 CATGCTCGCAAGCTCTCAGGGGGCGCGCAGTGTCTCCAACTTTTCGCGCTCCTCTACA 420
QY 421 GCTCTCTCCACATGCAAGTCTGCTCGCAGCAGCTTCTCCACCTTTTCTCTTTCTT 480
Db 421 GCTCTCTCCACATGCAAGTCTGCTCGCAGCAGCTTCTCCACCTTTTCTCTTTCTT 480
QY 481 TCTCTTTTCTTGCCCATCTTTGGTATTTTCAAAATGTCCCCCTTCAAAATGATAATCA 540
Db 481 TCTCTTTTCTTGCCCATCTTTGGTATTTTCAAAATGTCCCCCTTCAAAATGATAATCA 540
QY 541 CCAAAACTCATGAGCTGCTAGTATATAAATCTTAATCTTAAGTTTGGTGTATTTGAG 600
Db 541 CCAAAACTCATGAGCTGCTAGTATATAAATCTTAATCTTAAGTTTGGTGTATTTGAG 600
QY 601 TGGATTTTGTGCAAAAGTTGGTGTAGAAATAGGAGTTAAGGACCGCCAAACAAGATCCC 660
Db 601 TGGATTTTGTGCAAAAGTTGGTGTAGAAATAGGAGTTAAGGACCGCCAAACAAGATCCC 660
QY 661 CCACACTTAGCCCTTGTCTATCTCGAGTAAAGTTCAAGGACTAAGGTGGAACATCTCC 720
Db 661 CCACACTTAGCCCTTGTCTATCTCGAGTAAAGTTCAAGGACTAAGGTGGAACATCTCC 720
QY 721 TCAAAATGCTACGATGCTGCATATAGTTATTTCCAAAGCCTCACCTATACATGTGAATTT 780
Db 721 TCAAAATGCTACGATGCTGCATATAGTTATTTCCAAAGCCTCACCTATACATGTGAATTT 780
QY 781 GAAGTGTCTACACGCGCATCTTGGGTGTTGAGAAATGGAATGGAACAGATCAGATCCAGTCAT 840
Db 781 GAAGTGTCTACACGCGCATCTTGGGTGTTGAGAAATGGAATGGAACAGATCAGATCCAGTCAT 840
QY 841 CTTTACCTCTTGTCTAGATAACTTGGGTGTTTGTAAAGTTTCAAAATTTAAACATAG 900
Db 841 CTTTACCTCTTGTCTAGATAACTTGGGTGTTTGTAAAGTTTCAAAATTTAAACATAG 900
QY 901 TCTTGTCTCTCAAAATGATCTCTCATATAGCTCAATGTGTATGGTTTCTCACCAGGCCAA 960
Db 901 TCTTGTCTCTCAAAATGATCTCTCATATAGCTCAATGTGTATGGTTTCTCACCAGGCCAA 960
QY 961 TGTTTTGGCTCTTTTTCATCTTCTAATAATTTCTTTTGTGAGCTTAGGGTAGGGAAT 1020
Db 961 TGTTTTGGCTCTTTTTCATCTTCTAATAATTTCTTTTGTGAGCTTAGGGTAGGGAAT 1020
QY 1021 GAAAAGGAGCATACTTGCATTCATGCTTACTAAGTCAAAACCAAAATCTGAGGAGAA 1080
Db 1021 GAAAAGGAGCATACTTGCATTCATGCTTACTAAGTCAAAACCAAAATCTGAGGAGAA 1080
QY 1081 GCAAGTCAATCAATCTGATCAAGATGTGCAAGTGTGATATGTGGATTAAGATAAATC 1140
Db 1081 GCAAGTCAATCAATCTGATCAAGATGTGCAAGTGTGATATGTGGATTAAGATAAATC 1140
QY 1141 CTGTTTATTCATGCTCTCTCTTAATAAACTTTAGAGGGCATGGCAATCTTTGATGGG 1200
Db 1141 CTGTTTATTCATGCTCTCTCTTAATAAACTTTAGAGGGCATGGCAATCTTTGATGGG 1200

QY 1201 CTTTCATGAGCTCATCGTATGTCTAAGCATGAGCTCATCATTTATATATAGCATGTGAT 1260
Db 1201 CTTTCATGAGCTCATCGTATGTCTAAGCATGAGCTCATCATTTATATATAGCATGTGAT 1260
QY 1261 ACCAAAATTTACTCTTTTGGAGCATGTTTATATTTTAGGAGGACGTTTACCTGTTAGGTA 1320
Db 1261 ACCAAAATTTACTCTTTTGGAGCATGTTTATATTTTAGGAGGACGTTTACCTGTTAGGTA 1320
QY 1321 AATCTGAACGCTTAATAAATTCGGCTAAGCAAAAATAATTTATCACTGTTGATTCTAACAAT 1380
Db 1321 AATCTGAACGCTTAATAAATTCGGCTAAGCAAAAATAATTTATCACTGTTGATTCTAACAAT 1380
QY 1381 TTGATGATGCAACAATTTGATGAGTGACTGACAAATGATTTAGAGCTTTAAAGGAGATT 1440
Db 1381 TTGATGATGCAACAATTTGATGAGTGACTGACAAATGATTTAAAGGAGATT 1440
QY 1441 GAGAAGGATAAATCTACAATAAATAATGTAAGAAAGAACATTTCAAAGTGTGAGATCTGG 1500
Db 1441 GAGAAGGATAAATCTACAATAAATAATGTAAGAAAGAACATTTCAAAGTGTGAGATCTGG 1500
QY 1501 TGTGGAAGACTATTTTGGCTCTTGGGGGTAAAGACAACAAGTTTAGTAGTGCGCTCAA 1560
Db 1501 TGTGGAAGACTATTTTGGCTCTTGGGGGTAAAGACAACAAGTTTAGTAGTGCGCTCAA 1560
QY 1561 AATTGGGAGGCGCCATGCAAGATTTGTTAAAGTAAATTTGTTTGGATTTGACGGAGGCAATTC 1620
Db 1561 AATTGGGAGGCGCCATGCAAGATTTGTTAAAGTAAATTTGTTTGGATTTGACGGAGGCAATTC 1620
QY 1621 AAGGTGATCATCTACTAGAGCTCTCAATGGGAGGTGCTCGAAGACATATTAACCATGTG 1680
Db 1621 AAGGTGATCATCTACTAGAGCTCTCAATGGGAGGTGCTCGAAGACATATTAACCATGTG 1680
QY 1681 TATGCAAGATGTTTAGCTAGTAACTGACTGATAGTAAAGCATCTCAATGGGGCAAG 1740
Db 1681 TATGCAAGATGTTTAGCTAGTAACTGACTGATAGTAAAGCATCTCAATGGGGCAAG 1740
QY 1741 ACATATTAACCTAAGCGCGCTGTTTGTGCAAGTTTCGAGTAGGATATAGAGATCTCTCT 1800
Db 1741 ACATATTAACCTAAGCGCGCTGTTTGTGCAAGTTTCGAGTAGGATATAGAGATCTCTCT 1800
QY 1801 GCGAGTTGTAAACGATCTCCAATGGGGCAAGACATCTCACTATATATAGTGAAGGGC 1860
Db 1801 GCGAGTTGTAAACGATCTCCAATGGGGCAAGACATCTCACTATATATAGTGAAGGGC 1860
QY 1861 AGTACCTGATTCAGAAATCAATCAATCAAGCAACAATATAATTTTAAATTTTATTTCAA 1920
Db 1861 AGTACCTGATTCAGAAATCAATCAATCAAGCAACAATATAATTTTAAATTTTATTTCAA 1920
QY 1921 CCAATTTTTCCTTTTCCAAACCTTAATTAAGTTTTCCTTTTGGCTCTAGACAAAAT 1980
Db 1921 CCAATTTTTCCTTTTCCAAACCTTAATTAAGTTTTCCTTTTGGCTCTAGACAAAAT 1980
QY 1981 GACGTGTTCCGGGTATCTGCTGATTAAGAAACAACCCCTAGGTGACCTGTCCCGATAGA 2040
Db 1981 GACGTGTTCCGGGTATCTGCTGATTAAGAAACAACCCCTAGGTGACCTGTCCCGATAGA 2040
QY 2041 GTCCCACTGGGTAGGCATTCATAGGATTCGTTATTTCTCTGCAAAAAGCGATTAAAGC 2100
Db 2041 GTCCCACTGGGTAGGCATTCATAGGATTCGTTATTTCTCTGCAAAAAGCGATTAAAGC 2100
QY 2101 TGGCTTTTAAACTGGCTAGGCGGATTCGTGGGCTTCACTACCAAGGTGATTTTTCATGT 2160
Db 2101 TGGCTTTTAAACTGGCTAGGCGGATTCGTGGGCTTCACTACCAAGGTGATTTTTCATGT 2160
QY 2161 GATCCGTGATTTCTAGCACTTTTGTATGTAACCCAACTTAAGTGCACAATATATAATAT 2220
Db 2161 GATCCGTGATTTCTAGCACTTTTGTATGTAACCCAACTTAAGTGCACAATATATAATAT 2220
QY 2221 GCTACTTCGAGGATGTTATCAGACACAACCTCTTAATCTACGGAAGCCTTAAGTTTAGTT 2280
Db 2221 GCTACTTCGAGGATGTTATCAGACACAACCTCTTAATCTACGGAAGCCTTAAGTTTAGTT 2280

2281 TGCTCGGAGACAAGCAATTGTGGCCAGTCACTATAGCTACGTACAGAGGTAGTGGAGCA 2340
Db
2281 TGCTCGGAGACAAGCAATTGTGGCCAGTCACTATAGCTACGTACAGAGGTAGTGGAGCA 2340
Qy
2341 GTTGGCTGCTGGATTGAAACAGGTGGATCGTATCAGATATATGCAATTCATGGACA 2400
Db
2341 GTTGGCTGCTGGATTGAAACAGGTGGATCGTATCAGATATATGCAATTCATGGACA 2400
Qy
2401 GTAATGTGGTACAGTAACTTCGCAAAACAATAAATCTGTCAAAATTTATTTAGTGCATC 2460
Db
2401 GTAATGTGGTACAGTAACTTCGCAAAACAATAAATCTGTCAAAATTTATTTAGTGCATC 2460
Qy
2461 CTCTGACGTAATGCTTCTACGTCAGAGATTGATTTCCGAGGGCGCTGCACCCATCAC 2520
Db
2461 CTCTGACGTAATGCTTCTACGTCAGAGATTGATTTCCGAGGGCGCTGCACCCATCAC 2520
Qy
2521 TAATGACGGTCTTTACCCATCATATGACGACATTTGTTCAATCATGCTATCACTGTGCT 2580
Db
2521 TAATGACGGTCTTTACCCATCATATGACGACATTTGTTCAATCATGCTATCACTGTGCT 2580
Qy
2581 CCGTGTCCATGCACTGCAGCCCTCTATAAATCTGGCATCCCTCCCGTTACAGATCAC 2640
Db
2581 CCGTGTCCATGCACTGCAGCCCTCTATAAATCTGGCATCCCTCCCGTTACAGATCAC 2640
Qy
2641 ACAACACAAGCAAGAAATAAAGCGTAGCTGCCATAACTAGTACACCATGG 2690
Db
2641 ACAACACAAGCAAGAAATAAAGCGTAGCTGCCATAACTAGTACACCATGG 2690

RESULT 2

US-10-845-059-10
; Sequence 10, Application US/10845059
; Publication No. US2005005323A1
; GENERAL INFORMATION:
; APPLICANT: Birch, Robert George
; APPLICANT: Wu, Luguang
; TITLE OF INVENTION: METHOD FOR INCREASING PRODUCT YIELD
; FILE REFERENCE: 900145.401
; CURRENT APPLICATION NUMBER: US/10/845, 059
; PRIOR FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: AU2003902253
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 987
; TYPE: DNA
; ORGANISM: Saccharum sp.
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(987)
; OTHER INFORMATION: DNA sequence of promoter p67B

US-10-845-059-10

Query Match 31.4%; Score 843.8; DB 18; Length 987;
Best Local Similarity 91.7%; Pred. No. 6.3e-220;
Matches 947; Conservative 0; Mismatches 32; Indels 54; Gaps 3;
Qy 1639 GAGCTCTCAATGGAGGTGCTCGAAGACATATTAACCATGTGTATGGCAAGATGTTTAGC 1698
Db 1 GAGCTCTCAATGGAGGTGCTCGAAGACATATTAACCATGTGTATGGCAAGATGTTTAGC 60
Qy 1699 TAGTAACCTACCTAGTGTAAACGATCTCCATGGGGCAAGACATATTAACCTAAGGCCA 1758
Db 61 TAGTAGCTGACTGATAGTGTAAACGATCTCCATGGGGCAAGACATATTAACCTAAGGCCA 120
Qy 1759 GGCTGCTGTTTTGCAAGTTCAGTAGATATAGAGATTCTCGTGGAGTTGTAAACGATCT 1818
Db 121 GGCTGCTGTTTTGCAAGTTCAGTAGATATAGAGATTCTCGTGGAGTTGTAAACGATCT 180
Qy 1819 CCAATGGGGCAAGACATCTTACCTATATATAGTGAAGGGGAGTACGTGATGAGAATC 1878
Db 181 CCAATGGGGCAAGACATCTTACCTATATATAGTGAAGGGGAGTACGTGATGAGAATC 240

Qy 1879 AATCAATCAAGCACAAATATAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 1938
Db 241 AACCAATCAAGCACAAATATAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 299
Qy 1939 CCAACCCCTAAATATATAGTGTCT 1998
Db 300 CCAACCCCTAAATATATAGTGTCT 358
Qy 1999 TGCTGAATTAAGAACCAACCTAGGTGCACCTGTCCGATAGAGTCCCACTCGGTAGGCA 2058
Db 359 -----CCTGGGTAGGCA 370
Qy 2059 TTCTATAGGATTCGTGTATTTTCTCTCAAAAAAGCGATTAAAGCTGGCTTCTAAAACTGGCT 2118
Db 371 TTCTATAGGATTCGTGTATTTTCTCTCAAAAAAGCGATTAAAGCTGGCTTCTAAAACTGGCT 430
Qy 2119 AGGCCGGAATTCGTGGCTTCT 2178
Db 431 AGGCCGGAATTCGTGGCTTCT 490
Qy 2179 CTTTGTCTATGTAAACCAAACTTAAGTCGACAACTATAAATATGCTACTTGCAGGATGTTA 2238
Db 491 CTTTGTCTATGTAAACCAAACTTAAGTCGACAACTATAAATATGCTACTTGCAGGATGTTA 550
Qy 2239 TCACGACACAACTCTCTAATCTACGGAAGCCTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAA 2298
Db 551 TCATGACACAACTCTCTAATCTACGGAAGCCTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAA 610
Qy 2299 TGTGCCAGTCACTATAGCTACGTACAGAGGTAGTGGAGCAGTTGCGTGTGGATTGA 2358
Db 611 TGTGCCAGTCACTATAGCTACGTACAGAGGTAGTGGAGCAGTTGCGTGTGGATTGA 670
Qy 2359 AAACAGGTGGATCGTATCAGATATTTATGCAATTCATGACAGCAAGTAAATGCTGACAGTAA 2418
Db 671 AAACAGGTGGATCGTATCAGATATTTATGCAATTCATGACAGCAAGTAAATGCTGACAGTAA 726
Qy 2419 CTTTCCAAACAAATAAAATCTGTCAAAATTTATAGTGCACTCTCTGACGTAATGCTTTC 2478
Db 727 CTTTCCAAACAAATAAAATCTGTCAAAATTTATAGTGCACTCTCTGACGTAATGCTTTC 786
Qy 2479 TACGTACAGAGGATTTGATTCGAGGGCGCTGCACCCATCACTAATGACGGTCTTTTACCC 2538
Db 787 TACGTACAGAGGATTTGATTCGAGGGCGCTGCACCCATCACTAATGACGGTCTTTTACCC 846
Qy 2539 ATCATCATGGACCATTTGTCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 2598
Db 847 ATCATCATGGACCATTTGTCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 906
Qy 2599 CCTCTATAAATACCTGGCATCCCTCCCGGTTTACAGATCACACAAACCAAGCAAGAAAT 2658
Db 907 CCTCTATAAATACCTGGCATCCCTCCCGGTTTACAGATCACACCAACCAAGCAAGAAAT 966
Qy 2659 AAACGGTAGCTGC 2671
Db 967 AAACGGTAGCTGC 979

RESULT 3

US-10-437-963-1638
; Sequence 1638, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
NUMBER OF SEQ ID NOS: 204966
SEQUENCE ID NO: 1638
LENGTH: 4830
TYPE: DNA
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_101485C.1
US-10-437-963-1638

Query Match 7.2%; Score 192.4; DB 18; Length 4830;
Best Local Similarity 67.2%; Pred. No. 5.2e-41;
Matches 328; Conservative 0; Mismatches 136; Indels 24; Gaps 3;
QY 1232 GAGCTCATCATTTATATAGCATGGTGATACCAAAATTAATCTCTTTTGAGCATGTTTATA 1291
DB 4343 GGGCTCATCGTATCTAAGCATGGTGCAACAAAGGTAAACCAATTTGAATTAGTATATG 4402
QY 1292 TTTAGAGAGCGTTTACCTGTGAGGTAAATCTGAACGCTAATAATCGGCTAAGCAAA 1351
DB 4403 GGCAGAGAGCCATTTTGCATGTTGAGGTGAATCTGACGCGCTTAGATTGGCCAGACAA 4462
QY 1352 ATAAATTTATCACTGTTGATTTCTAACAATTTGATGATGACAAATATTGATGAGTGACTG 1411
DB 4463 ATAAATTTGCGCAGTAGATTAATCACTTAATGATGACGAATAGATGAGGTAGTG 4522
QY 1412 ACAATGATTTGAAGGCTTTAAAGGAGATTGAGAAGGATAAAT-----CT 1455
DB 4523 ATGAAAGATTGAAGCTTTTGAGAGAGATTGAGAAGATTAATTTGAGAGTAGCTAAAGCTT 4582
QY 1456 ACAATAAAATGTAAGAGAAAGCAATTCAAAGTGTGAGATCTGGTGTGGAAGACTATTT 1515
DB 4583 ACAATAAAAGGTGAAGGAAATCTGTTCAAAATTTGAGATTAGTGTGGAAGAAACAATTC 4642
QY 1516 TGCTCTTGGGGTAAAGACAAACAGTTTGTAGTGGCTCAAAATTTGGGAGGCCCA 1575
DB 4643 TGCCTGTTTGGTTCAAAAGATTAATAATTTGGAAATGTTCTCAAAATTTGGGAGGCCCA 4701
QY 1576 TCAAGATTGTTAAAGTAAT-----TGTGTTGATTGACGAGGCAATTTCAAGGTGAT 1628
DB 4702 TATAGAATTTGTAAGATAGTCCCGGGAATTTCTATTGTCGCAAGTCTACGAGGATAT 4761
QY 1629 CATCTACCTAGAGCTCTCAATGGAGGTGCTCGAAGACATATTACCCATGTGTATGGCAA 1688
DB 4762 AAGTTGCTAGAGCTTTTAATGGAGATATTTAAAGAGTACTATCCAAAGTGTGTCGCAA 4821
QY 1689 GATGTTTA 1696
DB 4822 GATGCTTA 4829

RESULT 4
US-10-425-115-54556
Sequence 54556, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 54556
LENGTH: 4920
TYPE: DNA
ORGANISM: Zea mays
FEATURE:

OTHER INFORMATION: Clone ID: MRT4577_149756C.1
US-10-425-115-54556

Query Match 6.0%; Score 160.4; DB 18; Length 4920;
Best Local Similarity 63.1%; Pred. No. 3.1e-32;
Matches 308; Conservative 0; Mismatches 156; Indels 24; Gaps 3;
QY 1232 GAGCTCATCATTTATATAGCATGGTGATACCAAAATTAATCTCTTTTGAGCATGTTTATA 1291
DB 4433 GGGCTCATAGAATATCTAAACATAGTGTCTACTAAAGTATCTCTTTTGAGCTTGTCTATG 4492
QY 1292 TTTAGAGGAGCGTTTACCTGTTGAGTAAATCTGAACCTAATAATCGGCTAAGCAAA 1351
DB 4493 GGCAGAGAGCATGTTTACCTGTGGAATTAAGTTTGAATGCTGTGAGTTTGCAGACAA 4552
QY 1352 ATAAATTTATCACTGTTGATTTCTAACAATTTGATGATGGAACAATATTGATGAGTGACTG 1411
DB 4553 ATGATCTAATCTGCTACTGATTATTAATTTCAATGATGATAATATTGATGAGTGACCG 4612
QY 1412 ACAATGATTTGAAGGCTTTAAAGGAGATTGAGAAGTAA-----ATCT 1455
DB 4613 ACAAAAGGATGATAGCTTTGGGAGCAATAGAAAAGGACAAAGATCATAGTACGAGGCT 4672
QY 1456 ACAATAAAATGTAAGAAAGCAATTCAAAGTGTGAGATCTGGTGTGGAAGACTATTT 1515
DB 4673 ACACAGAGAGTCAAGCAAAATCATTTCAAGTAGGAGACTTGGTGTGGAAGACCATTC 4732
QY 1516 TGCCTCTTGGGGTAAAGACAAACAGTTTGTAGTGGCTCAAAATTTGGGAGGCCCA 1575
DB 4733 TGCCTCTAAGGAAT-AAAGACCGAAAGTTTCGGGAAATGTCGCAAGCTGGGAGGCTCT 4791
QY 1576 TCAAGATTGTTAAAGTAATTTGTTTGA-----TTGACGAGGCAATTTCAAGGTGAT 1628
DB 4792 TATAAAGTGAACACAGGTGATGTCGTAAATGCTATTGCTACAAACATTAACAAGCGAA 4851
QY 1629 CATCTACCTAGAGCTCTCAATGGAGGTGCTCGAAGACATATTACCCATGTGTATGGCAA 1688
DB 4852 GATTTGCTTAAAGCTTTGATGAGGCTTCCCTCAAGCAGTATCATCCAGTATGTCGCAA 4911
QY 1689 GATGTTTA 1696
DB 4912 AATGCTTA 4919

RESULT 5

US-10-425-115-159798
Sequence 159798, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 159798
LENGTH: 7593
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_77309C.1
US-10-425-115-159798

Query Match 6.0%; Score 160.4; DB 18; Length 7593;
Best Local Similarity 63.1%; Pred. No. 4e-32;
Matches 308; Conservative 0; Mismatches 156; Indels 24; Gaps 3;
QY 1232 GAGCTCATCATTTATATAGCATGGTGATACCAAAATTAATCTCTTTTGAGCATGTTTATA 1291

Db 7106 GGGCTCATAGAAATATCTAAGACATAGTGTCTACTAAGGTATCTCTCTTTGAGCTTGTCTATG 7165
QY 1292 TTTAGGAGACGTTTACCTGTTGAGTAAATCTGAACGCTAATAAATCGGCTAAGCAA 1351
Db 7166 GGCAGGACGTTTACCTGTGGAATTAAGTTTGNATCTGTCTCAGTTTGGCAGACAA 7225
QY 1352 ATAAATTTACCTGTTGATTTCAACAATTTGATGATGACAAATATTTGATGAGGTGACTG 1411
Db 7226 ATGATCTAACTACTACTGATTTATTAATTTCAATGATGACAAATATTTGATGAGGTGACCG 7285
QY 1412 ACAATGATTTAAGGCTTTAAAGGAGATTTGAGAGGATAA-----ATCT 1455
Db 7286 ACAAGAGGATGATAGCTTTTGGGAGCAATAGAAAAGGACAAAGATCATGCTAGCTAGGGCCT 7345
QY 1456 ACAATAAAAATGTAAGAAAGAACATTCAAAGTGTGAGATCTGGTGTGGAAGACTATTT 1515
Db 7346 ACAACAGAGGTCAAAGCAAAATCATTCCAAGTAGGAGACTTTGGTGTGGAAGACCATTC 7405
QY 1516 TGCCTCTTGGGGTAAAGACAAAGTTTGTAGTGGCCTCAAAATTTGGAGGGCCCA 1575
Db 7406 TGCCTCTAAGGAAT-AAAGACCGAAAGTTTCGGMAATGTTGCGCAAGCTGGGAGGCTCT 7464
QY 1576 TCAAGATTTTAAAGTAATTTGTT-----TTGATTTGACGGAGGCATTTCAAGGTGAT 1628
Db 7465 TATAAGTGAACACAGGTGATGTTTGGTAAACGCTTATTTACTACAAACATTCACAGGCAAA 7524
QY 1629 CATCTACCTAGAGCTCTCAATGGGAGGTGCTCGAAGACATATTACCCATGTGTATGGCAA 1688
Db 7525 GATTTACCTAAGCTTTGAAATGGCGTTTCCCTCAACACAGTACCATCTCTAGTAGTGGCAA 7584
QY 1689 GATGTTTA 1696
Db 7585 GATGCTTA 7592

RESULT 6

US-10-425-115-54546.
; Sequence 54546, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 54546
; LENGTH: 6624
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_149747C.1
US-10-425-115-54546

Query Match 5.9%; Score 158.8; DB 18; Length 6624;
Best Local Similarity 62.9%; Pred. No. 1e-31;
Matches 307; Conservative 0; Mismatches 157; Indels 24; Gaps 3;
QY 1232 GAGCTCATGATTTATAGCATGTTGATACCAAAATTAATCTCTCTTTGAGCATGTTTATA 1291
Db 6137 GGGCTCATGAAATATCTAGCATAGTGTCTAAGTATCTCTCTTTGAGCTTGTCTATG 6196
QY 1292 TTTAGGAGACGTTTACCTGTTGAGTAAATCTGAACGCTAATAAATCGGCTAAGCAA 1351
Db 6197 GGCAGGACGTTTACCTGTGGAATTAAGTTTGAATGCTGTCAGTTTGGCAGACAA 6256
QY 1352 ATAAATTTACCTGTTGATTTCAACAATTTGATGATGACAAATATTTGATGAGGTGACTG 1411
Db 6257 ATGATCTAACTACTACTGATTTATTAATTTCAATGATGGAATAATATTTGATGAGGTGACCG 6316

QY 1412 ACAATGATTTGAAGCTTTTAAAGGAGATTTGAGAAAGATAA-----ATCT 1455
Db 6317 ACAAGAGGATGATAGCTTTTGGGAGCAATAGAAAAGGACAAAGATCATGCTAGCTAGGGCCT 6376
QY 1456 ACAATAAAAATGTAAGAAAGAACATTCAAAGTGTGAGATCTGGTGTGGAAGACTATTT 1515
Db 6377 ACAACAGAGGTCAAAGCAAAATCATTTCCAAGTAGGAGACTTTGGTGTGGAAGACCATTC 6436
QY 1516 TGCCTCTTGGGGTAAAGACAAAGTTTGTAGTGGCCTCAAAATTTGGAGGGCCCA 1575
Db 6437 TGCCTCTAAGGAAT-AAAGACCGAAAGTTTCGGAAATGTTGCGCAAGCTGGGAGGGTCT 6495
QY 1576 TCAAGATTTTAAAGTAATTTGTT-----TTGATTTGACGGAGGCATTTCAAGGTGAT 1628
Db 6496 TATAAGTGAACACAGGTGATGTTTGGTAAACGCTTATTTACTACAAACATTCACAGGCAAG 6555
QY 1629 CATCTACCTAGAGCTCTCAATGGGAGGTGCTCGAAGACATATTACCCATGTGTATGGCAA 1688
Db 6556 GATTTATCTAAGGCTTTGAAATGGCGTTTCCCTCAACACAGTACCATCTCTAGTAGTGGCAA 6615
QY 1689 GATGTTTA 1696
Db 6616 GATGCTTA 6623

RESULT 7

US-10-425-115-54557
; Sequence 54557, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 54557
; LENGTH: 3975
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_149757C.1
US-10-425-115-54557

Query Match 5.8%; Score 157.2; DB 18; Length 3975;
Best Local Similarity 62.7%; Pred. No. 2e-31;
Matches 306; Conservative 0; Mismatches 158; Indels 24; Gaps 3;
QY 1232 GAGCTCATGATTTATAGCATGTTGATACCAAAATTAATCTCTCTTTGAGCATGTTTATA 1291
Db 3488 GGGCTCACAGAATATCTAAACATAGTGTCTCTCTCTTTGAGCTTGTCTATG 3547
QY 1292 TTTAGGAGACGTTTACCTGTTGAGTAAATCTGAACGCTAATAAATCGGCTAAGCAA 1351
Db 3548 GGCAGGAGCAGTGTTCCTGTGGAATAAGTTTGAATGCTGTGAGTTCCGACAGACAA 3607
QY 1352 ATAAATTTACCTGTTGATTTCTAACAATTTGATGATGACAAATATTTGATGAGGTGACTG 1411
Db 3608 ATGATTTAACTCTCACTGATTTATCAATGCAATGATGATAATATTTGATGAGGTGACCG 3667
QY 1412 ACAATGATTTGAAGCTTTTAAAGGAGATTTGAGAGGATAA-----ATCT 1455
Db 3668 ACAAGAGGTTGATAGCTTTTGGGAGCAATAGAAAAGGACAAAGATCATCGTAGCCAGGCT 3727
QY 1456 ACAATAAATCTAAGAGAGAGCATTCAAAGTGTGAGATCTGGTGTGGAAGACTATTT 1515
Db 3728 ACAATAAGAGGTCAAGCAAAAGTCATTTCAAGTAGGGACCTTGGTGTGGAAGACCATCC 3787

QY 1516 TGCCTCTCGGGTAAAGACACAAAGTTTAGTAAGTGGCTCAAAATTTGGAGGGCCCA 1575
Db |||||
QY 3788 TGCCTCTAAGGAAT-AAAGACCGGAAGTTTGGAAGAAATGGTCGCAAGCTGGAGGGTCT 3846
Db |||||
QY 1576 TGAAGATTGTTAAAGTAAATTTGTTTGGG-----TTGACGAGGAGCATTTCAAGGTGAT 1628
Db |||||
QY 3847 TATAAGTAAACAGGTAATATCTGTAAGCCCTATTACTACAACATTAACAAGCAAG 3906
Db |||||
QY 1629 CATCTACCTAGAGCTCTCAATGGGAGGTGCTCGAAGACATATTACCCATGTGTATGCAAA 1688
Db |||||
QY 3907 GATTACCCCAAGCTTTGAATGGCGTTTCTCAAGCAGTACCCTCTAGTATGTGCAAA 3966
Db |||||
QY 1689 GATGTTTA 1696
Db |||||
QY 3967 GATGCTTA 3974
Db |||||

RESULT 8
US-10-437-963-4747
; Sequence 4747, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 4747
; LENGTH: 1031
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_11601C.1
US-10-437-963-4747

Query Match 5.8%; Score 156.8; DB 18; Length 1031;
Best Local Similarity 62.8%; Pred. No. 1.1e-31;
Matches 305; Conservative 0; Mismatches 157; Indels 24; Gaps 3;
QY 1232 GAGCTCATCTTTATATAAGCATGGTGATACCAAAATTTACTCTTTTGAGCATGTTTATA 1291
Db |||||
QY 547 GAGCTATAGATATCTAAGCATGGTGATCAAGTTTACTCTTTTGAGTTAATCTATT 506
Db |||||
QY 1292 TTTAGAGGACGTTTACCTGTTGAGTAAATCTGAACGCTAATAAATCGGCTAAGCAAA 1351
Db |||||
QY 607 TGAAGAAGCGGTTTACTAGTAGAGGTAATCTTGGCTCTGTAGATATATCAAGCAAG 566
Db |||||
QY 1352 ATAAATTTACCTGTTGATCTTCAAAATTTGATGATGACATATTTGATGAGTGACTG 1411
Db |||||
QY 667 ATGATTTATCGGCCGAAGAATATAGACGTTGATGGGAGCAATCTTGATGATGTGACCG 726
Db |||||
QY 1412 ACAAAATGATTGAAGGCTTTAAAGGAGATTGAGAAGGATATAA-----TCT 1455
Db |||||
QY 727 ATAAAGCTCTAAAGGCTTTAGAGGATAGAAAAGGAAAAGAGAGTAGCAAGGCT 786
Db |||||
QY 1456 ACATATAAATGTAAGAGAAAGCATTCAAAGTGTGAGATCTGGTGTGAAGACTATTTT 1515
Db |||||
QY 787 ATATAAAGAGAGTGAGAGAAAATCAATTTCAAGTAGGAAAATTTGGTTTGAAGACGATTT 846
Db |||||
QY 1516 TGCCTCTCGGGTAAAGACACAAAGTTTAGTAAGTGGCTCAAAATTTGGAGGGCCCA 1575
Db |||||
QY 847 TTCCT-TTGGGTTCTAGATTAAGACTTCGGTAAGTGGTTGCTAGATGGGAAGTCTT 905
Db |||||
QY 1576 TGAAGATTGTTAAAGTAAATTTGTTTGGG-----TTGACGAGGAGCATTTCAAGGTGAT 1628
Db |||||

Db 906 TATCGGATATGGAATCGTTTCAGGGAATGCGTATTTCTTGAAACTTTGCAAGGAGAA 955
QY 1629 CATCTACCTAGAGCTCTCAATGGGAGGTGCTCGAAGACATATTACCCATGTGTATGCAAA 1688
Db |||||
QY 966 TGATTTTAAAGAGCAATCAATGGGAATATTTGAAGAAATACTTCCCAAGCGTCTGCAAA 1025
Db |||||
QY 1689 GATGTT 1694
Db |||||
QY 1026 GATGCT 1031
Db |||||

RESULT 9
US-10-437-963-85940
; Sequence 85940, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 85940
; LENGTH: 1126
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_85028C.1
US-10-437-963-85940

Query Match 5.8%; Score 156.2; DB 18; Length 1126;
Best Local Similarity 62.6%; Pred. No. 1.7e-31;
Matches 305; Conservative 0; Mismatches 158; Indels 24; Gaps 3;
QY 1232 GAGCTCATCTTTATATAAGCATGGTGATACCAAAATTTACTCTTTTGAGCATGTTTATA 1291
Db |||||
QY 639 GGGCTCATCGGATATCTAAGCATGGTGCGACTCAAGTAACACCGTAGAGCTAGTCTATG 698
Db |||||
QY 1292 TTTAGAGGACGCTTTTACCTGTTGAGGTAATCTGAACGCTAATAAATCGGCTAAGCAAA 1351
Db |||||
QY 699 GGCAGAAGCGCGTTTACTTTGAGGTTAATTTACAAGCTCTTAGAATAGCCAGACAAA 758
Db |||||
QY 1352 ATAAATTTACCTGTTGATCTTCAAAATTTGATGATGACATATTTGATGAGTGACTG 1411
Db |||||
QY 759 ATGGTTTGTGCGCTGTAGATTATAGGAATTTAATGATGGACAGGATAGATGACGCTCCAG 818
Db |||||
QY 1412 ACAAAATGATTGAAGGCTTTAAAGGAGATTGAGAAGGATATAA-----CT 1455
Db |||||
QY 819 AGAAAGATTGAAGGCTTTGAGACATTTTGAAGAGAAAATTAAGAGTAGTAGGCTT 878
Db |||||
QY 1456 ACAATAAATAATGTAAGAGAAAAGCATTCAAAGTGTGAGATCTGGTGTGGAAGACTATTT 1515
Db |||||
QY 879 ATAAAGAGAGGTGAGAGAAAATCATTTTCAGATAGGTGAGTTGGTATGGAAGAACGATTT 938
Db |||||
QY 1516 TGCCTCTCGGGTAAAGACACAAAGTTTAGTAAGTGGCTCAAAATTTGGAGGGCCCA 1575
Db |||||
QY 939 TCCCATTTGGAATAGGAG-TAGCAAAATTTGCAAGTGGTCCCGAGCTGGGAAGTCTT 997
Db |||||
QY 1576 TCAAGATTGTTAAAGTAAATTTGTTTGGATTGA-----?-CGAGGCAATTTCAAGGTGAT 1628
Db |||||
QY 998 TATAAGTTGTAGAGATTGTTCTCGGAATTCATATTTTGTGAGTCTTTGAGGGGAT 1057
Db |||||
QY 1629 CATCTACCTAGAGCTCTCAATGGGAGGTGCTCGAAGACATATTACCCATGTGTATGCAAA 1688
Db |||||

Db 1058 AAGTTGCTAAAGCTTTGAATGGAAGATATTTGAAAAAATATTTATCCAGTATATGGTAG 1117
QY 1689 GATGTTT 1695
Db 1118 GGAGCTT 1124

RESULT 10
US-10-425-115-159738
; Sequence 159738, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 159738
; LENGTH: 5913
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_77253C.1
US-10-425-115-159738

Query Match 5.8%; Score 155.6; DB 18; Length 5913;
Best Local Similarity 62.5%; Pred. No. 7.1e-31;
Matches 305; Conservative 0; Mismatches 159; Indels 24; Gaps 3;
QY 1232 GAGCTCATCTTTATATAGCATGTGTGATACCAAAATTAATCTCTTTGAGCATGTTTATA 1291
Db 5426 GGGCGCATAGATATCTAAACATAGGCTACTAAAGTATCTCTTTGAGCTTGTCTATG 5485
QY 1292 TTTAGGAGGACGTTTACCTGTTGAGGTAAATCTGAACGCTAATAATCGGCTAAGCAAA 1351
Db 5486 GCGAGGAACAGTGTTCCTGTGGAAATAAGTTTGAATGCTATCAGGTTCCGCCAGACAAA 5545
QY 1352 ATAAATTTACCTGTGTGATTTCTAACTTTGATGATGGACATAATTTGATGAGGTGACTG 1411
Db 5546 ATGATCTAATCTGCTACCGATTATATGATTCAATGATGATGATAATATTGATGAGTGACTG 5605
QY 1412 ACAATGATTTGAAGGCTTTAAAGGAGATTGAGAGGATAAA-----TCT 1455
Db 5606 ACAAAAGGATGATAGCTTTTGGGAGCTATAGAAAAAGACAAAATCATGTGTAGCCAGGCGCT 5665
QY 1456 ACAATAAAAAATGTAAGAAGAACGATTCAAAGTGTGAGATCTGGTGTGGAAGACTATTT 1515
Db 5666 ACAACAAAAGGTCAAAGCAAAATCTTTCAAGTAGGGGACCTGGTGTGGAAGACCAATTC 5725
QY 1516 TGCCTCTTGGGGTAAAAGACAACTTTAGTAAAGTGGCCCTCAAAATTTGGGAGGCGCCA 1575
Db 5726 TGCCTCTAAGGAGTAA-AGACCGGAGTTCGGGAATGTCCACCAAGCTGGGAGGCGCT 5784
QY 1576 TCGAAGATTGTTAAAGTAATGTTTGGG-----TTGACGAGGAGCATTTCAAGGTGAT 1628
Db 5785 TACAAAGTAAAAACAGGTGATGTTTGGCAACGCGCTATTTCGTACAAACATTAACAAGGCAAG 5844
QY 1629 CATCTACCTAGAGCTCTCAATGGGAGTGTCTCGAGACATATTTACCATGTGTATGGCAA 1688
Db 5845 GACTTGCCTAAGGCTTTGAATGGGCGCTTTCCTCAAAACAGTACATCTCTAGTATGTGGCAG 5904
QY 1689 GATGTTTA 1696
Db 5905 GATGCCTA 5912

RESULT 11
US-10-437-963-25937

; Sequence 25937, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 25937
; LENGTH: 4014
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_30775C.1
US-10-437-963-25937

Query Match 5.8%; Score 155; DB 18; Length 4014;
Best Local Similarity 62.4%; Pred. No. 8.2e-31;
Matches 305; Conservative 0; Mismatches 160; Indels 24; Gaps 3;
QY 1232 GAGCTCATCTTTATATAGCATGTGTGATACCAAAATTAATCTCTTTGAGCATGTTTATA 1291
Db 3527 GGGCGCATAGGATATCTAAACATAGGCTGCGCACTAAAGTCACTCTTTGAGTTGGTTATG 3586
QY 1292 TTTAGGAGGACGTTTACCTGTTGAGGTAAATCTGAACGCTAATAATCGGCTAAGCAAA 1351
Db 3587 GTCAAAGAACGCGTTTACCGGTGGAGGTAAATCTTGGTTCTCTTCGTTATATCAAGCAAG 3646
QY 1352 ATAAATTTACCTGTGTGATTTCTAACTTTGATGATGACCAATATTGATGAGGTGACTG 1411
Db 3647 ATGATTTGTCAGTGAAGATTACAGACATTCATGGGAGACAACTTTGATGAGTCAATCG 3706
QY 1412 ACAATGATTTGAAGGCTTTAAAGGAGATTGAGAGGATAAA-----ATCT 1455
Db 3707 ACAACGCTTTGAAGGCAATTTGGAGGAGATAGAAAAGGAGAAAGAGGGGTGGCCAGGCGT 3766
QY 1456 ACAATAAAAAATGTAAGAAGAACGATTCAAAGTGTGAGATCTGGTGTGGAAGACTATTT 1515
Db 3767 ACAACAAAGAGGTGAAAGCAAAATTTGTTTCAAGTTGGAGACTTGGTTTGGAAAGCAATTT 3826
QY 1516 TGCCTCTTGGGGTAAAAGACAACTTTAGTAAAGTGGCCCTCAAAATTTGGGAGGCGCCA 1575
Db 3827 TGCCT- TTGGGTACTCGATCCAAAGGAGTTTCGTTAGTGTCTCTTAGTTGGGAGGTCTCT 3885
QY 1576 TCGAAGATTGTTAAAGTAATGTTTGGG-----TTGACGAGGAGCATTTCAAGGTGAT 1628
Db 3886 TATCGAGTGTGGCGCATTTTCGAGTGAACGCATACTTTTGGAGACACTTTCAAGGGGAG 3945
QY 1629 CATCTACCTAGAGCTCTCAATGGGAGTGTCTCGAGACATATTTACCATGTGTATGGCAA 1688
Db 3946 CGTTTTTCAGCGAGCAATCAACGGGAAATACTTTGAAGAAATACTTCCCGAGTGTGTTGCAA 4005
QY 1689 GATGTTTAG 1697
Db 4006 GACGCTTAG 4014

RESULT 12
US-10-437-963-26057/c
; Sequence 26057, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.

US-10-425-115-54555

Query Match 5.7%; Score 154; DB 18; Length 5460;
Best Local Similarity 62.3%; Pred. No. 1.9e-30;
Matches 304; Conservative 0; Mismatches 160; Indels 24; Gaps 3;

QY 1232 GAGCTCATCATTTATATAGCATGGTGATACCAAAATTACTCTTTTGAGCATGTTTATA 1291
DB 4973 GGGCTCATAGAAATATCAACATAGGGCTACTAAAGTATCTCCTTTTGAGCTTGCTATG 5032

QY 1292 TTTAGGAGACGCTTTTACCTGTTGAGTAAATCTGAACCTAATAAATCGGCTAAGCAAA 1351
DB 5033 GGCAGGAACGAGTGTGGCTGTGGAATAAGTTTGAATGCTGTCAAGTTTGCCAGACAAA 5092

QY 1352 ATAAATTTATCAGCTGTTGATTTCTAACAATTTGATGATGACAAATATTGATGAGTGACTG 1411
DB 5093 ATGATCTAACTGTTACTAATTAATTAATTCATGATGGAATAATATTGATGAGTGACCG 5152

QY 1412 ACAATATGATTTAAGGCTTTTAAAGGAGATTGAGAAGGATAA-----ATCT 1455
DB 5153 ACAAGAGGATGATAGCTTTTGGGAAGCAATAGAAAAGGACAAGATCATGTGTAGCCAGGCT 5212

QY 1456 ACAATATAAATGTAAGAAGAAAGCAATTCAAAGTGTGAGATCTGGTGTGGAAGACTATTT 1515
DB 5213 ACAACAAGAGGCTCAAAACAAAATCATTTTCAAGTAGGGGACCTGGTGTGGAAGACCATTC 5272

QY 1516 TGCCTCTTGGGGTAAAGACACAAAGTTTGTAGTGGCTCAAAATTTGGAGGGGCCCA 1575
DB 5273 TGCATCTAAGGAAT-AAAGACCGAAGTTTGGGAATGTTCGCAAGCTGGGAGGGTCT 5331

QY 1576 TGCAGAGTTGTTAAAGTAATGTTTGGGA-----TTGACGGAGGCATTTCAAGGTGAT 1628
DB 5332 TATAAGTAAACACAGGTGATGCTGCAACGGCTATTTGCTACAAACATTACAAGGCAAG 5391

QY 1629 CATCTACCTAGAGCTCTCAATGGAGGTGCTCGAAGACATATTACCCATGTGTATGGCAA 1688
DB 5392 GATTTGCTTAAGGCTTTGAATGGGCGTTTCTCCTCAACAGTACCATCTCTAGTATGTGCAA 5451

QY 1689 GATGTTTA 1696
DB 5452 GATGCTTA 5459

RESULT 15

US-10-425-115-159747
; Sequence 159747, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 159747
; LENGTH: 7170
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_77261C.1
US-10-425-115-159747

Query Match 5.7%; Score 154; DB 18; Length 7170;
Best Local Similarity 62.3%; Pred. No. 2.2e-30;
Matches 304; Conservative 0; Mismatches 160; Indels 24; Gaps 3;

QY 1232 GAGCTCATCATTTATATAGCATGGTGATACCAAAATTACTCTTTTGAGCATGTTTATA 1291
DB 6683 GGGCTCATAGAAATATCAACATAGTGTCTACTAAAGTATCTCCTTTTGAGCTTGCTATG 6742

QY 1292 TTTAGGAGACGCTTTTACCTGTTGAGTAAATCTGAACGCTAATAAATCGGCTAAGCAAA 1351
DB 6743 GGCAGGAAGCAGTGTCTCTGTGGAATAAGTTTGAATGCTCTCAGGTTTGCAGACAAA 6802

QY 1352 ATAAATTTATCAGCTGTTGATTTCTAACAATTTGATGATGGAACAATATTGATGAGTGACTG 1411
DB 6803 ATGATCTAACTGCTACTGATTTATTAATTTCAATGATGGAATAATATTGATGAGTGACCG 6862

QY 1412 ACAATATGATTTAAGGCTTTAAGGAGATTGAGAAGGATAA-----ATCT 1455
DB 6863 ACAAGAGGATGATAGCTTTGGGAGCAATAGAAAAGGACAAGATCATGTGTAGCCAGGCT 6922

QY 1456 ACAATATAAATGTAAGAAGAAAGCAATTCAAAGTGTGAGATCTGGTGTGGAAGACTATTT 1515
DB 6923 ACAACAAGAGGCTCAAGCAAAATCATTCCAAGTAGGAGACTTGGTGTGGAAGACCATTC 6982

QY 1516 TGCCTCTTGGGGTAAAGACAAAGTTTGTAGTGGCTCAAAATTTGGAGGGGCCCA 1575
DB 6983 TACCTCTAAGGAAT-AAAGATCGAAAATTCGGGAAAATGGTCGCAAGCTGGGAGGGTCT 7041

QY 1576 TGCAGAGTTGTTAAAGTA-----TTGTTTGGATTGACGGAGGCATTTCAAGGTGAT 1628
DB 7042 TATAAGTAAACACAGGTGATGCTGGTAATGCTTATTTACTACAAACATTACAAGGCAAG 7101

QY 1629 CATCTACCTAGAGCTCTCAATGGAGGTGCTCGAAGACATATTACCCATGTGTATGGCAA 1688
DB 7102 GATTTACCTAAGGCTTTGAATGGGCGTTTCTCCTCAACAGTACCATCTCTAGTATGTGCAA 7161

QY 1689 GATGTTTA 1696
DB 7162 GATGCTTA 7169

Search completed: May 9, 2005, 05:14:09
Job time : 1604 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2005, 21:19:18 ; Search time 11588 Seconds
(without alignments)
11248.234 Million cell updates/sec

Title: US-10-751-550-1

Perfect score: 2690

Sequence: 1 tctagataatagactact.....ccataactagatcacaccatgg 2690

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1030.2	38.3	1963	8 AJ626722	Saccharum officinarum partial dbp gene for putative dirigent protein.
2	192.4	7.2	107491	8 AP004613	Oryza sat
3	192.4	7.2	113514	8 AP004767	Oryza sat
4	188.6	7.0	144596	8 AP004574	Oryza sat
5	188.6	7.0	165766	8 AP004706	Oryza sat
6	182.8	6.8	123369	2 OSJN00028	Oryza sat
7	182.8	6.8	176627	8 AP006062	Oryza sat
8	181.4	6.7	139201	8 AP002970	Oryza sat
9	181.4	6.7	150022	8 AP003213	Oryza sat
10	175.8	6.5	136858	8 AC134345	Oryza sat
11	175.8	6.5	144294	8 AP004663	Oryza sat
12	175.8	6.5	149435	8 AC136521	Oryza sat
13	175.8	6.5	154416	8 AP005389	Oryza sat
14	173.6	6.5	127117	8 OSJN00040	Oryza sat
15	172	6.4	150137	8 AC135226	Oryza sat
16	171.2	6.4	198867	8 AP005325	Oryza sat
17	170.8	6.3	130521	8 OSJN00021	Oryza sat
18	170.8	6.3	145766	8 OSJN00287	Oryza sat
19	170.4	6.3	102515	8 H0806H05	Oryza sat

20	170.2	6.3	121591	8 OSJN00120	Oryza sat
21	169.6	6.3	137779	8 AP005843	Oryza sat
22	169.6	6.3	149581	2 AP005484	Oryza sat
23	169.6	6.3	177529	8 AP005305	Oryza sat
24	169.2	6.3	128574	8 OSJN00199	Oryza sat
25	169.2	6.3	152702	8 AP005162	Oryza sat
26	169.2	6.3	160462	8 OSJN00172	Oryza sat
27	169.2	6.3	184235	8 AP005166	Oryza sat
28	169.2	6.3	187972	8 AY360390	Oryza sat
29	166.8	6.2	138715	8 OSJN00190	Oryza sat
30	166.2	6.2	137297	8 AC090714	Oryza sat
31	164.8	6.1	171790	8 AP006460	Oryza sat
32	162.8	6.1	148373	8 AC134348	Oryza sat
33	162.6	6.0	4239	8 AK069351	Oryza sat
34	162.6	6.0	117075	8 AP004180	Oryza sat
35	162.6	6.0	148060	8 AP003334	Oryza sat
36	161.8	6.0	120640	8 OSJN00113	Oryza sat
37	161.8	6.0	131457	2 AC144718	Oryza sat
38	161.6	6.0	136214	8 AC134517	Oryza sat
39	161.6	6.0	169500	8 AC115686	Oryza sat
40	161.6	6.0	300029	8 AE017077	Oryza sat
41	161	6.0	176150	8 AC135864	Oryza sat
42	160.4	6.0	145528	2 AP005507	Oryza sat
43	160.4	6.0	158749	8 AY360394	Oryza sat
44	160.4	6.0	160541	8 AP004458	Oryza sat
45	160.4	6.0	188058	8 AP005819	Oryza sat

ALIGNMENTS

RESULT 1	AJ626722	1963 bp	DNA	linear	PLN 06-FEB-2004
LOCUS	Saccharum officinarum partial dbp gene for putative dirigent protein.				
DEFINITION	AJ626722				
ACCESSION	AJ626722.1	GI:42454401			
VERSION	dirigent protein; dbp gene.				
KEYWORDS	Saccharum officinarum				
SOURCE	Saccharum officinarum				
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum complex.				
REFERENCE	1				
AUTHORS	Abraham, T.G., Rogbeer, O., Groenewald, S., Groenewald, J.-H. and Botha, F.C.				
TITLE	Characterisation of culm-specific promoter elements in sugarcane				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 1963)				
AUTHORS	Abraham, T.G.				
TITLE	Direct Submission				
JOURNAL	Submitted (03-FEB-2004) Abrahama T.G., Botany, Institute for Plant Biotechnology, Private Bag X1, Matieland, 7602, SOUTH AFRICA				
FEATURES	Location/Qualifiers				
source	1..1963				
	/organism="Saccharum officinarum"				
	/mol_type="genomic DNA"				
	/variety="N19"				
	/isolation_source="Welgevallen experimental farm"				
	/db_xref="taxon:4547"				
	/clone="PGUS-DBP"				
	/tissue_type="culm"				
	/country="South Africa:Stellenbosch"				
gene	1..1715				
	/gene="dppb"				
promoter	1..1151				
	/gene="dppb"				
	/function="guiding protein"				
CDS	1152..1715				
	/gene="dppb"				
	/codon_start=1				

/product="putative dirigent protein"
/protein_id="CAF25234.1"
/db_xref="GI:42454402"
/translation="WASRSVLAVSLALALFAVAPASFPALDERLHLSLYLNQVSGN
GLNQAVVPEGLPGFEGNNAIQDPVFNNGESDQVVGRAQGIQFKPSERNDOAWYTT
LTVIFVTSLSGLQMMGYIFQDQWISFGGTGQLTWARGVNVHKKVVRQINGRMKY
INIAFYTPFLGASSCRINPRDLFDA"

ORIGIN

Query Match 38.3%; Score 1030.2; DB 8; Length 1963;
Best Local Similarity 94.9%; Pred. No. 4.3e-246;
Matches 1136; Conservative 0; Mismatches 13; Indels 48; Gaps 5;
1 1490 GTGAGATCTGGTGTGAAGACTATTTTGGCTCTTGGGGTAAAGACAACAGTTAGTA 1549
Db 1 GTAAGATCTGGTGTGAAGACTATTTTGGCTCTTGGGGTAAAGACAACAGTTAGTA 60
1550 AGTGGCTCAAAATGGGAGGGCCCATGCAAGATCTTAAAGTAAATGTTTGGATTGAC 1609
Db 61 AGTGGCTCAAAATGGGAGGGCCCATGCAAGATCTTAAAGTAAATGTTTGGATTGAC 120
1610 GGAGGCATTCAAGGTGATCACTACCTAGAGCTCTCAATGGGAGGTGCTGGAAGACATA 1669
Db 121 GGAGGCATTCAAGGTGATCACTACCTAGAGCTCTCAATGGGAGGTGCTGGAAGACATA 180
1670 TTACCCATGTATGCAAGATGTTTGTAGCTAGTACCTGATGATGATGATGATGATGATG 1729
Db 181 TTACCCATGTATGCAAGATGTTTGTAGCTAGTACCTGATGATGATGATGATGATGATG 240
1730 AATGGGGCAAGACATATTACCTAAGCCAGGCTGGTGTGTTTGAAGTTCAGTAGGATATA 1789
Db 241 AATGGGGCAAGACATATTACCTAAGCCAGGCTGGTGTGTTTGAAGTTCAGTAGGATATA 300
1790 GAGATTCTCGTGGAGTGTGAAACGATCTTCAATGGGCAAGACATCTTACCTCATATAT 1849
Db 301 GAGATTCTCGTGGAGTGTGAAACGATCTTCAATGGGCAAGACATCTTACCTCATATAT 360
1850 AGTGAAGGGCAGTACCTGATGAGATCAATCAATCAAGCAACATATATATTAATTAAT 1909
Db 361 AGTGAAGGGCAGTACCTGATGAGATCAATCAATCAAGCAACATATATATTAATTAAT 420
1910 TTTTATTTCAAAACCAATTTTTCCTTTTCCAAACCTTAATATATATATATATATATAT 1969
Db 421 TTTTATTTCAAAACCAATTTTTCCTTTTCCAAACCTTAATATATATATATATATATAT 479
1970 TAGGACAAATGACGTGTCGGGTATCTGCTGTAATTAAGAACACCTAGTGCACCT 2029
Db 480 TAGGACAAATGACGTGTCGGGTATCTGCTGTAATTAAGAACACCTAGTGCACCT 539
2030 GTCCCGATAGTCCCACTGGTGGTATTCATAGGG-ATTGCTGATTTTCTGCG-AAA 2087
Db 540 GTCCCGATAGTCCCACTGGTGGTATTCATAGGGATTCGGGTATTTCTGCGAAA 599
2088 AAAGCATTAAAGTGGCTTTCTAAACCTGAGCGCGATTCGTGGCCTTCACTACCAAG 2147
Db 600 AAAGCATTAAAGTGGCTTTCTAAACCTGAGCGCGATTCGTGGCCTTCACTACCAAG 659
2148 GTGATTTTCAATGATCCGTGCAATTCAGCACTTGTGTAATGAACCCAACTTAAGTGA 2207
Db 660 GTGATTTTCAATGATCCGTGCAATTCAGCACTTGTGTAATGAACCCAACTTAAGTGA 719
2208 CAACTATAATATGCTACTTGGAGGATGTTATCAGCACCACTCTTAATCTACGGAAGC 2267
Db 720 CAACTATAATATGCTACTTGGAGGATGTTATCAGCACCACTCTTAATCTACGGAAGC 778
2268 CTAAGTTTATGTTTGTTCGAGACAAGCAATTTGGCCAGTCACTATAGCTAGCTCAGAG 2327
Db 779 CTAAGTTTATGTTTGTTCGAGACAAGCAATTTGGCCAGTCACTATAGCTAGCTCAGAG 838
2328 GGTAGTGGGAGCAGTTGCGTGTGATTTGAATGAACAGGTGGATCGTATCAGATATATGC 2387
Db 839 GGTAGTGGGAGCAGTTGCGTGTGATTTGAATGAACAGGTGGATCGTATCAGATATATGC 869

QY 2388 ATTCACATGGACAGTAATAATGTGTACGTAACTTCGCAAAACAATAAAATCTGTCAAAAT 2447
Db 870 -----AAATGTGTACGTAACTTCGCAAAACAATAAAATCTGTCAAAAT 914
QY 2448 TATTAGTGCATCTCTGACGTAATAATGTTTACGTGACAGGATTTGATTCGAGGCGG 2507
Db 915 TATTAGTGCATCTCTGACGTAATAATGTTTACGTGACAGGATTTGATTCGAGGCGG 974
QY 2508 CTGCAACCATCACTAATGACGGTCTTTACCCATCATGACCATTTGTTCATCCATG 2567
Db 975 CTGCAACCATCACTAATGACGGTCTTTACCCATCATGACCATTTGTTCATCCATG 1034
QY 2568 CTATCACTGTGCTCTGTCATGCACTGCAGCCCTCTATAAATACCTGCGATCCCTCC 2627
Db 1035 CTATCACTGTGCTCTGTCATGCACTGCAGCCCTCTATAAATACCTGCGATCCCTCC 1094
QY 2628 GTTCACAGATCACAAACACACAGCAAGCAATAAAGCGTAGCTGCCATACTAGTACA 2684
Db 1095 GTTCACAGATCACAAACACACAGCAAGCAATAAAGCGTAGCTGCCATACTAGTACA 1151
AP004613 107491 bp DNA linear PLN 31-AUG-2004
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,
BAC clone:OJ1123.G09.
AP004613 BR000010
AP004613.3 GI:21328156
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1
Sasaki,T., Matsumoto,T. and Yamamoto,K.
Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, BAC
clone:OJ1123.G09
Published Only in Database (2002)
2 (bases 1 to 107491)
Sasaki,T., Matsumoto,T. and Yamamoto,K.
Direct Submission
Submitted (10-JAN-2002) Takuji Sasaki, National Institute of
Agricultural Sciences, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
On Jun 6, 2002 this sequence version replaced gi:19773520.
Genes were predicted from the integrated results of the following:
GENSCAN1.0, BLASTX2.0, BLAST2.0 as well as SplicePredictor
(October 1998 version). The genomic sequence was searched against
NCBI Nonredundant Protein database, nr
(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at
RGP. Protein homologs of the coding regions were searched against
NCBI Nonredundant Protein database with BLASTP2.0. ESTs represent
the identified cDNA sequences using BLASTN 2.0 with the
corresponding DBJ accession no. and RGP clone ID.
A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative-', and '-like protein'. A gene without
significant homology to any protein but with EST homology (covering
almost the entire length of partial sequence) is classified as an
'unknown' protein. A gene predicted with a gene prediction program
is classified as a 'hypothetical' protein.
The orientation of the sequence is from -21M13 to M13rev of the BAC
clone. This sequence of OJ1123.G09 clone has an overlap with
OJ1460.H08 (DBJ: AP004767) at the position 86,928 to 107,491 of 3',
end. The nucleotide sequence of this BAC clone was generated by
combining Monsanto and RGP-Japan sequencing data. Detailed
information on overlap and assembly quality together with
annotation of this entry is available at
http://rgp.dna.affrc.go.jp/GenomesSeq.html.
Location/Qualifiers

FEATURES

```
source
1..107491
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="1"
/clone="OJ1123_G09"
/complement(3..1367)
/gene="OJ1123_G09.1"
/complement(3..1367)
/misc_feature
/gene="OJ1123_G09.1"
/note="probably inactive due to stop codon(s) in CDS
pseudogene, gag-pol precursor"
/pseudo
complement(1601..2944)
/gene="OJ1123_G09.2"
complement(1601..2944)
/gene="OJ1123_G09.2"
/note="hypothetical protein
similar to Oryza sativa chromosome 10, OSJNBa0011A24.17"
/codon_start=1
/protein_id="BAC00734.1"
/db_xref="GI:21328157"
/translation="MSYVDPRTSPVQPLXYGVYPPSTVACQVGGIGAGCSTMTS
SFTDSVNTLNGVSTTLLVMTQGEIVFPVTVILISAGPPMAGNENAVARTRGD
SMSKNPAAEIGTSTSEPEKDPAAKPCLSNRHEPMRITSEVTRSWCPJHKTRKH
TLQAWVFLNHAEIRACKECGIGTSPTRDVHCPIHKTKNHDLSEKCVLSAMKTPP
PKVQSRIPRDEDEKQGNRPILDFVGVIDIDHPSVLHLLDYGSTSTSTPREVL
AIDVGVTRTNAEENQVAPTQHIRAINAILRPTDPDINDQQAIRWELRQESVT
NLDVGVFEAAATPEQPTDANGENPERRESHPPTPPRPTGCDLLDHLHNGREAR
HTRDNERSRHVSSRAIQAKIEAIAQTETVITITITMIANGAGRYWSRTP"
3436..4231
/gene="OJ1123_G09.3"
misc_feature
3436..4231
/gene="OJ1123_G09.3"
/note="probably inactive due to stop codon(s) in CDS
pseudogene, transposable element"
/pseudo
join(4645..5386,6323..6438)
/gene="OJ1123_G09.4"
join(4645..5386,6323..6438)
/gene="OJ1123_G09.4"
/note="hypothetical protein
similar to Oryza sativa chromosome 1, P0043B10.18"
/codon_start=1
/protein_id="BAC00735.1"
/db_xref="GI:21328158"
/translation="WHKQVMESSLETKLIDMPOQETNGDEAAPEADISMATLTM
PNGVERKATAPALPAARPPPEMTDKVMASSTANLAQLPTGTALAYQALSTFTNHG
OCYRNRLWLTAGLAVLTASSITFFSLTDSVVGKGLYVGMATPRGFVFNLSREEE
AOELSRKTLRELVRPLDIVHAFFTVAVFLTVAFSDVGLTKCFPPDAGNDTKELLKNL
PLGMAFMSTFVFLFPYKRGIGYDTTPRIIDGINKTNFFDTPTKGNFCSGVGEKR
ARLSTTHCG"
join(7843..8274,8373..8505,8786..8946)
/gene="OJ1123_G09.5"
join(7843..8274,8373..8505,8786..8946)
/gene="OJ1123_G09.5"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC00736.1"
/db_xref="GI:21328159"
/translation="WASGGMRSPATATGAGGGGQGRQRRRPKVTAVHARRTAVH
TGERRRGRGEPHVRWPERKAAVMIGDAKGTGGDGTIGERGEGLGGRKPAAT
KNGQEMERMEKLRSRARKFGAHLTSTSVHAPGSENFYKCSSTRILDRGSSCTS
IKVFINLIDVSRLEVTFCVCSVGHFVRGEVLQRYSVLLTTKFGNFSIGDGDKTEAT
RMEIVPESRSGSKRIGCDWDR"
join(12918..13037,13203..13586,14357..14887,15596..16339)
/gene="OJ1123_G09.6"
join(12918..13037,13203..13586,14357..14887,15596..16339)
/gene="OJ1123_G09.6"
/codon_start=1
/product="putative polypeptide"
/protein_id="BAC00737.1"

/db_xref="GI:21328160"
/translation="WPPKKVAKENVARREKESEAGVEMAAEAGEPSALGAEADVAASP
QQAATASTAHSTPTSKPKPTGCPVRIWCDSTRTCFVKCKSSRRHTKWPCKPSRPR
PPFSIROATSHPRSTKCRPKPRHKSCHQRTSPDSLRRPOYIRPLMYSHKFSRSL
PKYAPRHLPTKEQTHLLGIRQPRDSIKYMYCRFSQARQCVQDITFASINASAGLLE
GELTRKIANKEPQTVHEHLRIIDYARGEEDSKRQAIOAEDYKASVIAAQVQVQV
AEPTPLAVCQSQANQGPQSQPMTRKFRFTDSAGKAVMAVEVEALRKEFDAAQ
ATSHQQPVKKLKMPPAGVIVKGLQPSATS KGLAAINRAVHDVEAEAPHDRAKHAP
KPAPHGRVKEQVNDADSTFRVSLGGDMEREVENILEVLKKNIDIFTWGPDEVGVS
TDLIMHILAVKPDARLRKQLRKMSADQEAKEVQKQKAGVIOEIDHPWLANPV
LVRSNGKRMVCVDFTLNKAYPKDDFPLPIDQLVDSTAGCELSFLDAYSGYHQIH
MNPADIPKATFITPSTFCHSGCLSA"
join(21504..21802,21856..22234)
/gene="OJ1123_G09.7"
join(21504..21802,21856..22234)
/gene="OJ1123_G09.7"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC00738.1"
/db_xref="GI:21328161"
/translation="MTVGESMSTVKGRAPVDHPKRWRRKSWCELGENGVRTVLGCG
EGAIEVGSDATKPELVAPRKEVLTAKLRSSATAAARESERESNDSSFRVECGMGR
GGLESGBLVSNPVAFNPGCKICGRISGQPKRREEREGKWRERERTRVNPFPQAG
NGMAMVRRGQQRREVGGDRWVRMLASVCGAVCVQGRPSKRGADQADGRNRNG
NGHIITL"
complement(25948..26166)
/gene="OJ1123_G09.8"
complement(25948..26166)
/gene="OJ1123_G09.8"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC00739.1"
/db_xref="GI:21328162"
/translation="MPFLLVGKIKTKLDKNMPPGRFLRSSLVLPFGFCPKQKFCSP
26726..27024
/note="5' LTR"
27071..31705
/gene="OJ1123_G09.9"
27071..31705
/gene="OJ1123_G09.9"
/note="probably inactive due to stop codon(s) in CDS
pseudogene, gypsy/Ty-3 retroelement polypeptide"
/pseudo
31718..32012
/note="3' LTR"
complement(join(33051..33520,33601..33713,33765..33994))
/gene="OJ1123_G09.10"
complement(join(33051..33520,33601..33713,33765..33994))
/gene="OJ1123_G09.10"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC00740.1"
/db_xref="GI:21328163"
/translation="MAAARTKSGRRARPGCTMDLHRWSTGERSASQLRGGAWPDD
HRQDRDRRHGKQNGRRRGCGGSGPHRAPRPPATGTGRRVRAEHLDLAKMGEG
LGREEVRRGGRCVAGCGGCGVIGACGRRERWQVAGGSDLLNFAVKVIMWLC
FWGGVDAGVROGTVMQWABRGDGGSCGSRPEKVVVRVCGGYSMLKRWPRC
GTRRRSRWWHSAATANTMACDRSMLASGEALGVSRGERSGLGEGK"
complement(join(34860..34963,35445..35480,36798..36852,
36971..37123,38978..39065,39351..39376))
/gene="OJ1123_G09.11"
complement(join(34860..34963,35445..35480,36798..36852,
36971..37123,38978..39065,39351..39376))
/gene="OJ1123_G09.11"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC00741.1"
/db_xref="GI:21328164"
/translation="MVPSAVNLSPISGHKDARRSLNPPSKALSPMAQLQGOWPDL
RSKPLVIDSTVGTTRNSCLLFCYNQVILDFWSELCLFELDCILTLHSHRRPSQFA
SSVHDAQKHYSERKNGKLSLVHRGSRSTWTHGLGRPNSTITSLTIGA"
join(44553..44618,45251..45275,46869..46927,46990..47209,
```

```
47387..47664)
/gene="OJ1123_G09.12"
join(44553..44618,45251..45275,46869..46927,46990..47209,
47387..47664)

Query Match
Best Local Similarity 7.2%; Score 192.4; DB 8; Length 107491;
Matches 328; Conservative 0; Mismatches 136; Indels 24; Gaps 3;

OY 1232 GAGCTCATCTTATATAAGCATGGTGATACCAAAATTACTCTTTTGAGCATGTTTATA 1291
Db 100032 GGGCTCATCTATATCTAAGCATGGTGCAACAAAGGTAAACACCATTTGAATTAGTATATG 100091

OY 1292 TTTAGGAGGAGCTTTTACCTGTTGAGGTAAATCTGACGCTAATAATCGGCTAACCAA 1351
Db 100092 GGCAGAGGAGCAATTTTGCATGTTGAGGTGAATCTGGACGCCCTTAGATTGGCCAGACAA 100151

OY 1352 ATAAATTTATCACCTGTTGATCTTAAACAATTTGATGATGACATAATTTGATGAGTCACTG 1411
Db 100152 ATAAATTTGCGGAGTAGATTATCACACTTAATGATGACGGAATAGATGAGTTAGTG 100211

OY 1412 ACAATGATTAAGGCTTTAAAGGAGATTGAGAGGATAAAT-----CT 1455
Db 100212 ATGAAAGATTGAAAGCTTTGAGAGAGATTGAGAAAGATAAATTCAGAGTAGCTTAAAGCTT 100271

OY 1456 ACAATAAATGTAAGAAGCAATTCAGAGTGTGAGATCTGGTGTGGAAGACTATT 1515
Db 100272 ACATAAAGGTAAGGAAATATCTTCAAAATGGAGATTAGTGTGGAAGCAATTC 100331

OY 1516 TGCTCTTGGGGTAAAGACAACTAGTTTGTAGTGGCTCAAAATTTGGGAGGCCCA 1575
Db 100332 TGCTCTT-GGTTCAAGATAATAATTTGGAAATGTCCTCAAAATTTGGGAGGCCCA 100390

OY 1576 TCGAAGATTGTTAAAGTAAT-----TGTTTGGATTGACGAGGCAATTCAGGTGAT 1628
Db 100391 TATAGAATTGTAGAAATAGTCCCGGGAATTTCTATTGTCGCAAGTCTACGAGGATAT 100450

OY 1629 CATCTACCTAGAGCTCTCAATGGAGGTGCTCGAAGACATATTACCCATGTGTGGCAA 1688
Db 100451 AGTTGCCCTAGAGCTTTAATGGAGATATTATTAAGGTACTATCCAGTGTGTGGCAA 100510

OY 1689 GATGTTTA 1696
Db 100511 GATGCTTA 100518

RESULT 3
AP004767
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) linear PLN 31-AUG-2004
BAC clone: OJ1460_H08.
ACCESSION
AP004767 BA000010
VERSION
AP004767.2 GI:20086415
KEYWORDS
Oryza sativa (japonica cultivar-group)
SOURCE
Oryza sativa (japonica cultivar-group)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1
Sasaki.T., Matsumoto.T. and Yamamoto.K.
Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, BAC
clone: OJ1460_H08
Published Only in Database (2002)
2 (bases 1 to 113514)
Sasaki.T., Matsumoto.T. and Yamamoto.K.
Direct Submission
Submitted (20-FEB-2002) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail: tsasaki@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/,
Tel: 81-298-38-7441, Fax: 81-298-38-7468)
On Apr 8, 2002 this sequence version replaced gi:18844983.

COMMENT
```

Genes were predicted from the integrated results of the following:
GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor
(October 1998 version). The genomic sequence was searched against
NCBI NonRedundant Protein database, nr
(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at
RGP. Protein homologies of the coding regions were searched against
NCBI NonRedundant Protein database with BLASTP2.0. ESTs represent
the identified cDNA sequences using BLASTN 2.0 with the
corresponding DBJ accession no. and RGP clone ID.
A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative-' and '-like protein'. A gene without
significant homology to any protein but with EST homology (covering
almost the entire length of partial sequence) is classified as an
'unknown' protein. A gene predicted with a gene prediction program
is classified as a 'hypothetical' protein.
The orientation of the sequence is from -21M13 to M13rev of the BAC
clone. This sequence of OJ1460_H08 clone has an overlap with
OJ1123_G09 (DBJ: AP004613) clone at 5' end and with B1109A06
(DBJ: AP004610) at 3' end. The sequence was generated by combining
Monsanto and RGP-Japan sequencing data. Detailed information on
overlap and assembly quality together with annotation of this entry
is available at
http://rgp.dna.affrc.go.jp/GenomeSeq.html.

location/Qualifiers

1..113514
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="1"
/clone="OJ1460_H08"
complement(join(778..811,2823..3109))
/gene="OJ1460_H08.1"
complement(join(778..811,2823..3109))
/gene="OJ1460_H08.1"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB92937.1"
/db_xref="GI:20805271"
/translation="MSVGEADSDAGSGSATTTGGGANGPRFSTCGGARRRGRCG
MATPWPSSRRFCALAIVGDSQFIEWRSVQLGDFANKVGERSFGMACPRTSHYIS
KLFP"

5635..5969
/note="5' LTR"
6334..13592
/gene="OJ1460_H08.2"
6334..13592
/gene="OJ1460_H08.2"
/note="probably inactive due to frameshift(s) in CDS
probably inactive due to no initiation codon in CDS
probably inactive due to stop codon(s) in CDS
pseudogene, putative retrotransposon Cinfu1 prpol"
/pseudo
13902..14237
/note="3' LTR"
complement(join(15314..15491,19080..19200,20389..20524,
21054..21209))
/gene="OJ1460_H08.3"
complement(join(15314..15491,19080..19200,20389..20524,
21054..21209))
/gene="OJ1460_H08.3"
/note="hypothetical protein
similar to Arabidopsis thaliana chromosome 3, At3g23610"
/codon_start=1
/protein_id="BAB92938.1"
/db_xref="GI:20805272"
/translation="MSMFEAGSRSDREDDYEQOARVLMALMQGGAARYKAD
NIPCPVGLVIGSVGAAMKDKALSLNTHLIVARSINPAFAEFYKKEVLDDSP
DIDLAKHDECFESFIDEISGGVNLVHCFAGRSRSVTIIIVLMKKHQMSENALSL
VRSKRPQVAPNEGFSQENFEKSMQGIQIAFTFK"
join(24874..25186,26967..27016)
/gene="OJ1460_H08.4"

FEATURES

source

gene

CDS

LTR

gene

misc_feature

LTR

gene

CDS

gene

```
CDS
join(24874..25186,26967..27016)
/gene="OJ1460_H08.4"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB92939.1"
/db_xref="GI:20805273"
/translation="MTLLVDDGDWVAAREARDAARPAARARRSGRGRDGRRLLEAV
EAAVFRVFRVVSAGMGKGRNSRACPCGGAAGGASGVARRWLGRRLSGGGGQWQW
RLALEISLVPQVMSRA"
complement(join(42130..42140,42359..42465,44531..44673))
/gene="OJ1460_H08.5"
complement(join(42130..42140,42359..42465,44531..44673))
/gene="OJ1460_H08.5"
/note="contains ESTs AU082610(C30277),AU062714(C30277)
unknown protein"
/codon_start=1
/protein_id="BAB92940.1"
/db_xref="GI:20805274"
/translation="WFSRHLRLVLLSISQPSPRRMVLKRVMSATVFTSGCRRHCRF
RLLRQKLWIFLVRVRIARALALFNNPTILLDEPTNLDMA"
53007..53915
/gene="OJ1460_H08.6"
53007..53915
/gene="OJ1460_H08.6"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB92941.1"
/db_xref="GI:20805275"
/translation="MAVTVAVRGFGPRTAEGSAPGADVAPTWRTRGHVARHAG
GRTDPLPEGRSDGGERTFGAHLKLPARPWRTAALPRGGGRRRRPTASGGGAT
ARAALAALWGLGGAARCKGRERERERERERERERERERERERERERERERERER
LANTVKVGREGEDGRGADGRRRPWRWRTARSGFORRNPTEGWIWPARRRTRWR
LFRGSVORLETASGGSGGGGEGDGRARAMEGTGDCCKTKKGGNGEPLRYRTGV
DRGRDRFPASVGRGGBEREVGFRN"
complement(54043..54476)
/note="3' LTR"
complement(54488..58879)
/gene="OJ1460_H08.7"
complement(54488..58879)
/gene="OJ1460_H08.7"
/note="probably inactive due to frameshift(s) in CDS
probably inactive due to stop codon(s) in CDS
pseudogene, gag-pol"
/pseudo
complement(join(62255..62275,62449..63042))
/gene="OJ1460_H08.8"
complement(join(62255..62275,62449..63042))
/gene="OJ1460_H08.8"
/note="hypothetical protein
similar to Oryza sativa chromosome 3, OSJNBb0031G04.25"
/codon_start=1
/protein_id="BAB92942.1"
/db_xref="GI:20805276"
/translation="WADHPNHRGDMDDFVAELARMTLVGVGNAPYSTVPPLSGE
LPHRVRLEVHCYLTGLTSLANVVEASGTAADHAYQEAALYMLMARLDHHSYIFCNAYR
LPHRRASGNDISTPRTGVGENTDTFGHMCVNRGMDRHSDLHKASKALNDGKLVRIL
ALKDLARLKENAQIKGLPAPGVRIKRTTPKSTTPFLIGII"
complement(66254..66702)
/note="5' LTR"
join(76774..76835,77289..77457)
/gene="OJ1460_H08.9"
join(76774..76835,77289..77457)
/gene="OJ1460_H08.9"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB92943.1"
/db_xref="GI:20805277"
/translation="MGVLTGRPRPPFAGSAGGANDFLAILALPKGVLITGNITPTG
VSTPAPARDESQEAIIIPGRGEACDRPPVAI"
78396..81540
/gene="OJ1460_H08.10"
78396..81540
/gene="OJ1460_H08.10"

CDS
join(24874..25186,26967..27016)
/gene="OJ1460_H08.4"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB92939.1"
/db_xref="GI:20805273"
/translation="MTLLVDDGDWVAAREARDAARPAARARRSGRGRDGRRLLEAV
EAAVFRVFRVVSAGMGKGRNSRACPCGGAAGGASGVARRWLGRRLSGGGGQWQW
RLALEISLVPQVMSRA"
complement(join(42130..42140,42359..42465,44531..44673))
/gene="OJ1460_H08.5"
complement(join(42130..42140,42359..42465,44531..44673))
/gene="OJ1460_H08.5"
/note="contains ESTs AU082610(C30277),AU062714(C30277)
unknown protein"
/codon_start=1
/protein_id="BAB92940.1"
/db_xref="GI:20805274"
/translation="WFSRHLRLVLLSISQPSPRRMVLKRVMSATVFTSGCRRHCRF
RLLRQKLWIFLVRVRIARALALFNNPTILLDEPTNLDMA"
53007..53915
/gene="OJ1460_H08.6"
53007..53915
/gene="OJ1460_H08.6"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB92941.1"
/db_xref="GI:20805275"
/translation="MAVTVAVRGFGPRTAEGSAPGADVAPTWRTRGHVARHAG
GRTDPLPEGRSDGGERTFGAHLKLPARPWRTAALPRGGGRRRRPTASGGGAT
ARAALAALWGLGGAARCKGRERERERERERERERERERERERERERERERERER
LANTVKVGREGEDGRGADGRRRPWRWRTARSGFORRNPTEGWIWPARRRTRWR
LFRGSVORLETASGGSGGGGEGDGRARAMEGTGDCCKTKKGGNGEPLRYRTGV
DRGRDRFPASVGRGGBEREVGFRN"
complement(54043..54476)
/note="3' LTR"
complement(54488..58879)
/gene="OJ1460_H08.7"
complement(54488..58879)
/gene="OJ1460_H08.7"
/note="probably inactive due to frameshift(s) in CDS
probably inactive due to stop codon(s) in CDS
pseudogene, gag-pol"
/pseudo
complement(join(62255..62275,62449..63042))
/gene="OJ1460_H08.8"
complement(join(62255..62275,62449..63042))
/gene="OJ1460_H08.8"
/note="hypothetical protein
similar to Oryza sativa chromosome 3, OSJNBb0031G04.25"
/codon_start=1
/protein_id="BAB92942.1"
/db_xref="GI:20805276"
/translation="WADHPNHRGDMDDFVAELARMTLVGVGNAPYSTVPPLSGE
LPHRVRLEVHCYLTGLTSLANVVEASGTAADHAYQEAALYMLMARLDHHSYIFCNAYR
LPHRRASGNDISTPRTGVGENTDTFGHMCVNRGMDRHSDLHKASKALNDGKLVRIL
ALKDLARLKENAQIKGLPAPGVRIKRTTPKSTTPFLIGII"
complement(66254..66702)
/note="5' LTR"
join(76774..76835,77289..77457)
/gene="OJ1460_H08.9"
join(76774..76835,77289..77457)
/gene="OJ1460_H08.9"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB92943.1"
/db_xref="GI:20805277"
/translation="MGVLTGRPRPPFAGSAGGANDFLAILALPKGVLITGNITPTG
VSTPAPARDESQEAIIIPGRGEACDRPPVAI"
78396..81540
/gene="OJ1460_H08.10"
78396..81540
/gene="OJ1460_H08.10"

LTR
Query Match 7.2%; Score 192.4; DB 8; Length 113514;
Best Local Similarity 67.2%; Pred. No. 2.7e-36;
Matches 328; Conservative 0; Mismatches 136; Indels 24; Gaps 3;

QY 1232 GAGCTCATCTTATATTAAGCATGTGTATACCAAAATTTACTCTCTTTTGTAGCATGTTTATA 1291
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
13105 GGGCTCATCTTATATTAAGCATGTGTATACCAAAATTTACTCTCTTTTGTAGCATGTTTATA 13164
QY 1292 TTTAGGAGAGCGTTTACTCTGTGAGTAAATCTCAACGCTAATAATCGGCTAAGCAAA 1351
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
13165 GCGAAGAGCGCAATTTGCATGTTGAGGTGAATCTGACGCCCTTAGATTGGCCAGACAA 13224
QY 1352 ATAAATTTATCACCTGTTGATTTCAACAATTTGATGATGACCAATATTGATGAGTGACTG 1411
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
13225 ATAAATTTGCGCAGTATATCACACTTAATGATGACGGAATAGATGAGTTAGT 13284
QY 1412 ACAATGATTAAGCGCTTTAAAGGAGATTGAGAGAGATAAAT-----CT 1455
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
13285 ATGAAAGATTGAAAGCTTTGAGAGAGATTGAGAAAGATAAATTTGAGAGTAGCTAAAGCTT 13344
QY 1456 ACAATAAAATGTAAGAGAAAGCATTCAAAGTGTGAGATCTGTTGTGGAAGACTATT 1515
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
13345 ACAATAAAAGGTGAAGGAAATAATCGTTTCAAATTTGAGATTAGTGTGGAATAAATTC 13404
QY 1516 TGCCTCTTGGGGTAAAAGACAAACAAGTTTACTAGTGGCCCTCAAATTTGGAGGGCCCA 1575
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
13405 TGCCTGTT-GGTTCAAAGATAATAAATTTGGAATGCTCTCCAAATTTGGAGGGCCA 13463
QY 1576 TCAAGATTGTTAAAGTAAT-----TGTGTTGATTGACGAGGCAATTTCAAGTGAT 1628
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
13464 TATAGAAATGTAGAAATAGTCCCGGAATCTTATTTTGTGCAAAAGTCTACGAGGATAT 13523
QY 1629 CATCTACCTAGAGCTCTCAATGGGAGGTCTCGAGACATATTTACCCATGTGTATGCGAA 1688
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
13524 AAGTTGCTTAGAGCTTTTAAATGGGAGATATTTTAAAGAGGTACTATCCAAAGTGTGTCGAA 13583
QY 1689 GATGTTTA 1696
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
13584 GATGCTTA 13591

misc_feature
RESULT 4
AP004574
```

LOCUS	AP004574	144596 bp	DNA	linear	PLN 03-DEC-2003
DEFINITION	Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 8, PAC clone:P0702G08.				
ACCESSION	AP004574				
VERSION	AP004574.3	GI:38636932			
KEYWORDS					
SOURCE	Oryza sativa (japonica cultivar-group)				
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.				
REFERENCE	1				
AUTHORS	Sasaki, T., Matsumoto, T. and Yamamoto, K.				
TITLE	Oryza sativa nipponbare (GA3) genomic DNA, chromosome 8, PAC clone:P0702G08				
JOURNAL	Published Only in Database (2001)				
REFERENCE	2 (bases 1 to 144596)				
AUTHORS	Sasaki, T., Matsumoto, T. and Yamamoto, K.				
TITLE	Direct Submission				
JOURNAL	Submitted (19-DEC-2001) Takuji Sasaki, National Institute of Agrobiological Sciences, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan				
COMMENT	<p>(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/, Tel:81-298-38-7441, Fax:81-298-38-7468)</p> <p>On Dec 2, 2003 this sequence version replaced gi:33235572. Genes were predicted from the integrated results of the following: GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH (http://www.softberry.com/), GeneMark.hmm (http://opal.biology.gatech.edu/GeneMark/), GlimmerM (http://www.tigr.org/tdb/glimmerm/glmr_form.html), RiceHMM (http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4 (http://globin.cse.psu.edu/html/docs/sim4.html), gap2 (http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The genomic sequence was searched against NCBI NonRedundant Protein database, nr (http://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RGP or DBJ. Protein homologies of the coding regions were searched against NCBI NonRedundant Protein database with BLASTP. ESTs represent the identified cDNA sequences using BLASTN with the corresponding DBJ accession no. and RGP clone ID. Full-length cDNAs represent the identified cDNA sequences using BLASTN with the corresponding DBJ accession no.</p> <p>A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with full-length cDNA or EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted by two or more gene prediction programs is classified as a 'hypothetical' protein according to IRGSP standard. A gene predicted by a single gene prediction program is also classified as a probable 'hypothetical' protein and is included as a miscellaneous feature of the sequence.</p> <p>The orientation of the sequence is from SP6 to T7 of the PAC clone. This sequence of P0702G08 clone has an overlap with P0683E12 (DBJ: AP004706) clone at 5' end and with OSJNBA0062G05 (DBJ: AP005491) at 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at http://rgp.dna.affrc.go.jp/GenomeSeq.html.</p>				
FEATURES	Location/Qualifiers				
Source	1. .144596 /organism="Oryza sativa (japonica cultivar-group)" /mol_type="genomic DNA" /cultivar="Nipponbare" /db_xref="taxon:39947" /chromosomes="8" /clone="P0702G08" complement (2836..3303) /gene="P0702G08.1" complement (2836..3303) /misc_feature /gene="P0702G08.1" /note="putative gypsy-type retrotransposon orf5" complement (4240..7583)				
gene					
misc_feature					
gene					

mRNA	gene="P0702G08.2" complement (join(4240..4693,4959..5060,5167..5244,5348..5441,5561..5739,5814..5930,6295..7048,7154..7585)) /gene="P0702G08.2" /note="supported by full-length cDNA(s): AK100576" complement (join(4562..4693,4959..5060,5167..5244,5348..5441,5561..5739,5814..5930,6295..6852)) /gene="P0702G08.2" /note="contains EST(s): AU085915 (E2817), C73047 (E2817) contains full-length cDNA(s): AK100576" /codon_start=1 /product="GTPase activator protein-like" /protein_id="BAD03195.1" /db_xref="GI:38636933" /translation="MFGTQVGVVDDVFNRRRTLTWNTPAASMTTTHTRISTPKKKGHANAHOILAKFEDLYGPMVGNVDDVNLVDVRRMRQGVVWVALEASGANWYLOQPRISNGSEGVISVTSKLKSLVNTITLRLIRKQVPPVLRPKVLVSGAAKKSSTVPETYYDLIRATEKTPATLQIDHDLPRFPFCHSWLNSEEGQASLRVLVGYSGFRFSEVGYCQGLNYVAALLLVLMKTEEDAFWMLVNLVNDYDYNLSCGCHVEORVFLDLLAKCPRIAAHLEAMGFDVSLVATEWFLCFKSTLPSETTLRVWDLVFNEGAKVLFHVALAIKMRKEDLLRIQHIGVIDILQITAAHLYDPDELLTFAFDKIGKSWTTNTITTKERKQETVYMAELDQIRRLSSLKMDG" complement (10029..10368) /gene="P0702G08.3" complement (join(<10029..10235,10267..>10368)) /gene="P0702G08.3" /note="start and end point are not identified" complement (join(10029..10235,10267..10368)) /gene="P0702G08.3" /note="predicted by GENSCAN etc." /codon_start=1 /product="hypothetical protein" /protein_id="BAD03196.1" /db_xref="GI:38636934" /translation="MLCKRTVLLPPSICTAALYRREETRRRSASFRRRFPASSTTARSPSAAATYPPPPAPPLRLQHHEDVRRYISSAATAPFPNRLAIARWTGKGGRGIGRIG
------	---

mrna

CDS

gene

mrna

CDS

gene

misc_feature

gene

misc_feature

gene

misc_feature

gene

mrna

CDS

by two or more gene prediction programs is classified as a 'hypothetical' protein according to IRGSP standard. A gene predicted by a single gene prediction program is also classified as a probable 'hypothetical' protein and is included as a miscellaneous feature of the sequence.

The orientation of the sequence is from 5' to 3' of the PAC clone. This sequence of P0683E12 clone has an overlap with P0689D06 (DBJ: AP004621) clone at 5' end and with P0702G08 (DBJ: AP004574) at 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at <http://fsg.dna.affrc.go.jp/GenomeSeq.html>.

FEATURES

Source

1. 165766
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="8"

gene

/clone="P0683E12"
complement(2914..4170)

mRNA

/gene="P0683E12.1"
complement(<2914..>4170)

CDS

/note="start and end point are not identified"
complement(2914..4170)
/gene="P0683E12.1"
/note="similar to Oryza sativa chromosome 4,
OSUNB0072M01.3"

/codon_start=1
/product="hypothetical protein"
/protein_id="BAD03373.1"
/db_xref="GI:38637119"

gene

/translation="MDEHAYLHPLDALEIVLHLSQSLADRAFRGVCQWAAVRDQ
WPTQPMPLAAGHCVDLSVHRVPLPSGDDVDGVCSSGLGNWLTALTPKRRWR
PHQVRPLLPFGASVOLPILTPAARFGGDDINVEKIVMSSAPSDGCVAAIIVM
GYSSTREIVWRQESCSAPAAAPSNVADAFVHGDLVYVDDKCSQLYVFNHVFSS
DGCGQELHVRLEMDLRTSRFARVLLECDRLIMADRHGKDAGVHEVRVVALE
RASCQDNEWSVPTRLDGHVFLGACCFRALPTGDRVKDGNVFLDDSAEITAVVT
VDRKPLERSALIRSMVPASNVLDTFRRGGGGGADRPAAPAAACWAGRRNOCFG
FGGLQLIVLLKSFVSLQEDGTHRQ"
complement(4870..5415)
/gene="P0683E12.2"
complement(join(<4870..4876,5264..>5415))
/gene="P0683E12.2"
/note="start and end point are not identified"
complement(join(4870..4876,5264..5415))
/gene="P0683E12.2"
/note="predicted by GeneMark.hmm etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD03374.1"
/db_xref="GI:38637120"

mRNA

/translation="MAAAELEDNASGDWRPSLGTARAAATATTVGGCGNGCAWGRNV
FIWVGVRN"
6094..13263
/gene="P0683E12.3"
join(<6094..6298,9294..9339,10053..10167,11333..11476,
11566..11892,12319..12512,13128..>13263)

CDS

/note="start and end point are not identified"
join(6094..6298,9294..9339,10053..10167,11333..11476,
11566..11892,12319..12512,13128..13263)
/gene="P0683E12.3"
/note="predicted by GeneMark.hmm etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD03375.1"
/db_xref="GI:38637121"

gene

/translation="WARARGLYGLASDNGNGRQRPQTGHTADNGRANSNGWRERGO
SSHTGGGREGAPLDWERTGAVGVDFCPAPACQIKGRALAGKQGYWVAATGS
TVVNHAPNLLTHFTFILLAQREGEAGARGKRLRLTRDGGGRRRRNKAAGR
LMGNLERRPAEAKAVEGEVDDGTARRRFGGGAARQGGGGLAGGVEAAADWSSAV
AAADWLGEVAAARWRRRIGWRRRLVAAVAAADWSGGDVEAAADWLVAARWRRRIGW

mRNA

/gene="P0683E12.4"
join(<14594..14973,15522..15790,15942..16055,16481..16584,
17601..>17717)
/gene="P0683E12.4"
/note="start and end point are not identified"
join(14594..14973,15522..15790,15942..16055,16481..16584,
17601..17717)
/gene="P0683E12.4"
/note="similar to Oryza sativa chromosome
10, OSUNBa0061H20.13"

CDS

/codon_start=1
/product="hypothetical protein"
/protein_id="BAD03376.1"
/db_xref="GI:38637122"

gene

/translation="MEVTITAIISVCVKYTSISFAFPSLHTMNHFDVHLPCASOPP
EHDHDALEHAPERREDIIVRPSSIPKQAMPRISSPRELQDASTLARNIKNT
ENTSFSSTLIRHURRRRLPLRWGPPGVPHLPLPARRAEATAAPTSPSPFG
HVNHARACASAPPTPPLAPAPEMAGIRKNSPASLSSPTTHVGRFPPLPATSGS
PEGCAVAVGFAMPSTPTSPSPFVAGALIPSRTAAPAFSPSPSPPAVLRRCR
NRAPSRAMAAGPARVGVMERGGPCDGLCRFRIHLGTRGDCPLPAGDGSET"
complement(join(18175..18279,18722..18772))
/gene="P0683E12.5"
complement(join(18175..18279,18722..18772))
/gene="P0683E12.5"
/note="hypothetical ORF
predicted by GeneMark.hmm
this category is not included in IRGSP standard"
23449..23936
/gene="P0683E12.6"
join(<23449..23467,23530..>23936)
/note="P0683E12.6"
/note="start and end point are not identified"
join(23449..23467,23530..23936)
/gene="P0683E12.6"
/note="predicted by GENSCAN etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD03377.1"
/db_xref="GI:38637123"

mRNA

/translation="MANOSIPFAARRGIPTNTSHLMKNRNSIKGNOHLVVYHIOHKN
KVEAPSSVEGLHCKTETEPWPWSSRDALREPLRELSTMGSRKMGWACGAMWSP
AVLPSELINIPRELFEVNVHGRSRRSAGRRQRRVETNS"
complement(24717..24878)
/gene="P0683E12.7"
complement(24717..24878)
/gene="P0683E12.7"
/note="hypothetical ORF
predicted by GlimmerM
this category is not included in IRGSP standard"
complement(25230..26573)
/gene="P0683E12.8"
complement(join(<25230..25246,26420..>26573))
/gene="P0683E12.8"
/note="start and end point are not identified"
complement(join(25230..25246,26420..26573))
/gene="P0683E12.8"
/note="predicted by GENSCAN etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD03378.1"
/db_xref="GI:38637124"

CDS

/translation="MWPGTLGDVGEGVEVSALLTFPEGGRRVTPILAPRAEAPRLAF
LLHYVGLQPCR"
27237..30542
/gene="P0683E12.9"
27237..30542
/gene="P0683E12.9"
/note="probably inactive due to including stop codon(s) in
CDS
pseudogene, retroelement"

AAAGSIDLSAIDHSACAGVSATPTPPSLAAGQLDAPPPPSAAAPRRPTATKPRR
RVCSPAAOFFGALTAGGAGQREWRRTAEWRRRTAAGVANGGEVVALTAVPSDA"
14594..17717
/gene="P0683E12.4"
join(<14594..14973,15522..15790,15942..16055,16481..16584,
17601..>17717)
/gene="P0683E12.4"
/note="start and end point are not identified"
join(14594..14973,15522..15790,15942..16055,16481..16584,
17601..17717)
/gene="P0683E12.4"
/note="similar to Oryza sativa chromosome
10, OSUNBa0061H20.13"

/codon_start=1
/product="hypothetical protein"
/protein_id="BAD03376.1"
/db_xref="GI:38637122"

/translation="MEVTITAIISVCVKYTSISFAFPSLHTMNHFDVHLPCASOPP
EHDHDALEHAPERREDIIVRPSSIPKQAMPRISSPRELQDASTLARNIKNT
ENTSFSSTLIRHURRRRLPLRWGPPGVPHLPLPARRAEATAAPTSPSPFG
HVNHARACASAPPTPPLAPAPEMAGIRKNSPASLSSPTTHVGRFPPLPATSGS
PEGCAVAVGFAMPSTPTSPSPFVAGALIPSRTAAPAFSPSPSPPAVLRRCR
NRAPSRAMAAGPARVGVMERGGPCDGLCRFRIHLGTRGDCPLPAGDGSET"
complement(join(18175..18279,18722..18772))
/gene="P0683E12.5"
complement(join(18175..18279,18722..18772))
/gene="P0683E12.5"
/note="hypothetical ORF
predicted by GeneMark.hmm
this category is not included in IRGSP standard"
23449..23936
/gene="P0683E12.6"
join(<23449..23467,23530..>23936)
/note="P0683E12.6"
/note="start and end point are not identified"
join(23449..23467,23530..23936)
/gene="P0683E12.6"
/note="predicted by GENSCAN etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD03377.1"
/db_xref="GI:38637123"

/translation="MANOSIPFAARRGIPTNTSHLMKNRNSIKGNOHLVVYHIOHKN
KVEAPSSVEGLHCKTETEPWPWSSRDALREPLRELSTMGSRKMGWACGAMWSP
AVLPSELINIPRELFEVNVHGRSRRSAGRRQRRVETNS"
complement(24717..24878)
/gene="P0683E12.7"
complement(24717..24878)
/gene="P0683E12.7"
/note="hypothetical ORF
predicted by GlimmerM
this category is not included in IRGSP standard"
complement(25230..26573)
/gene="P0683E12.8"
complement(join(<25230..25246,26420..>26573))
/gene="P0683E12.8"
/note="start and end point are not identified"
complement(join(25230..25246,26420..26573))
/gene="P0683E12.8"
/note="predicted by GENSCAN etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD03378.1"
/db_xref="GI:38637124"

/translation="MWPGTLGDVGEGVEVSALLTFPEGGRRVTPILAPRAEAPRLAF
LLHYVGLQPCR"
27237..30542
/gene="P0683E12.9"
27237..30542
/gene="P0683E12.9"
/note="probably inactive due to including stop codon(s) in
CDS
pseudogene, retroelement"

COMMENT	Tel:81-298-38-7441, Fax:81-298-38-7468)	
	On Feb 28, 2001 this sequence version replaced gi:11526597. Genes were predicted from the integrated results of the following: GENSCAN1.0, BLASTX2.0, BLASTX2.0 as well as SplicePredictor (October 1998 version). The genomic sequence was searched against NCBI NonRedundant protein database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RGP. Protein homologies of the coding regions were searched against NCBI NonRedundant Protein database using BLASTP2.0. ESTs represent the identified cDNA sequences using BLASTN 2.0 with the corresponding DDBJ accession no. and RGP clone ID. A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted with a gene prediction program is classified as a 'hypothetical' protein. The orientation of the sequence is from 996 to T7 of the PAC clone. Detailed information on overlap and assembly quality together with annotation of this entry is available at http://rgp.dna.affrc.go.jp/GenomeSeq.html.	
FEATURES	Location/Qualifiers	
	1..139201	
source	/organism="Oryza sativa (japonica cultivar-group)"	
	/mol_type="genomic DNA"	
gene	/cultivar="Nipponbare"	
	/db_xref="taxon:39947"	
CDS	/chromosome="1"	
	/clone="P0504D03"	
gene	join(2371..2591,2628..2925)	
	/genes="P0504D03.1"	
CDS	join(2371..2591,2628..2925)	
	/genes="P0504D03.1"	
gene	/note="hypothetical protein"	
	/codon_start=1	
CDS	/protein_id="BAB32955.1"	
	/db_xref="GI:13161360"	
gene	/translation="MAATVAVARLMAAREDIGEAKDDAVGDEADAAVDTARHVL GIESWTVAEPTFSAADLTITVEPAKTRGRSARQWRMAGGSEEDATAVGEEDNA EAVRRQCSGGCEGGRGGDEEAQWRVGRRTRRRQVREIRRRGGVAGRSRA VVAAGAEVFR"	
	complement(join(3966..4098,5324..5442,5530..5649, 5655..5711))	
CDS	/genes="P0504D03.2"	
	complement(join(3966..4098,5324..5442,5530..5649, 5655..5711))	
gene	/genes="P0504D03.2"	
	/note="hypothetical protein"	
CDS	/codon_start=1	
	/protein_id="BAB32956.1"	
gene	/db_xref="GI:13161361"	
	/translation="MQEUSLLDIELVDAPPHSDASLSPHPRRRSRFAPVLIDVE GVIPGMVEMRISRTRRCQPRALAEAMETHNGGRTGARRAVLAVNGRLNQ HVGPITLILIALITVHCDSPHSNMETHLSDG" join(14660..14794,15810..15929,16670..16752,16842..16953, 17025..17129,17244..17294,20044..20133,20217..20325, 20400..20548,20646..20699,22252..22455) /genes="P0504D03.3"	
CDS	join(14660..14794,15810..15929,16670..16752,16842..16953, 17025..17129,17244..17294,20044..20133,20217..20325, 20400..20548,20646..20699,22252..22455)	
	/genes="P0504D03.3"	
gene	/codon_start=1	
	/product="putative ARG1 protein (Altered Response to Gravity)"	
CDS	/protein_id="BAB32957.1"	
	/db_xref="GI:13161362"	
gene	/translation="WASSGMEGSPAPAMRDPYEVLSVPRDSIQEIKSAVRKLAK FHPDKNASPEALPFKEVAYSILSDPEKRRQYDTAGFEALNEGMDIEDLNLIG TVNTWPAALFSKLGVPDKITVSPNVLEAMSGTVTRPLPVGSSATGKVDKASAFYG VTISEQAGSIVRVTSQAQKFLFFQEGINGISYGLAQEDSQTKQVTSAGMYF LHFQVTRMDSVNALAMAKDPAAFKRLLEGLOPCEVSALKSGTHIFAVYGNDFKPA	
	complement(join(40009..40028,41743..41887))	
CDS	/genes="P0504D03.7"	
	complement(join(40009..40028,41743..41887))	
gene	/genes="P0504D03.7"	
	/note="hypothetical protein"	
CDS	/codon_start=1	
	/protein_id="BAB32961.1"	
gene	/db_xref="GI:13161366"	
	/translation="MANLPRVMGFGSHGDGVASNGTSPVPTITVGLPSCDQIGPEW KSALESDLHR"	
CDS	join(43590..43634,43673..43894)	
	/genes="P0504D03.8"	
gene	join(43590..43634,43673..43894)	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
gene	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	
CDS	complement(join(40009..40028,41743..41887))	
	/genes="P0504D03.8"	

```
/notes="hypothetical protein"
/codon_start=1
/protein_id="BAB32962.1"
/db_xref="GI:13161367"
/translation="MRERQRGPDVLLVEMTEGEGSVRPERVKADGEGDSAGAGGG
GRGEGGGGSGVGRGRRGRGGBESKLLSRAWIRPBKMRG"
gene
50026..50735
/misc_feature
50026..50735
/gene="P0504D03.9"
/gene="P0504D03.9"
/notes="probably inactive due to frameshift(s) in CDS
pseudogene
similar to Oryza sativa Sgcl"
/pseudo
/gene="P0504D03.10"
complement(join(51394..51430,52618..52850))
/gene="P0504D03.10"

Query Match 6.7%; Score 181.4; DB 8; Length 139201;
Best Local Similarity 63.2%; Pred. No. 1.6e-33;
Matches 367; Conservative 0; Mismatches 196; Indels 18; Gaps 5;

QY 486 TTCTTGGCCCATCTTTGGTATTTCACAAATGTCCTCCCTACAAATGATAATCACCAG 545
DB 72222 TTCTTGGCATGGTGGGTGTTTGTTCATAATTCACCTTTATACAAATATTCACCAAG 72163
QY 546 ACTCATGGAGCTTGTCTAGTATATAAATCTAAATCTAAGTTTGGTGTATTTTGGAG---- 600
DB 72162 ACTTATGGAATGGTTAGTATATACCACTACCTCTAAGTTTGNATCTATTTATATTCT 72103
QY 601 -----TGGATTTTGTGTGAAGTTGGTGTGTAATAATAGAGTTAAGGACCGCCCAAG 655
DB 72102 TTCAATTTGCTTTTATGCATGAATTCACAATCGATCTGCTACTTAAGGACTGTTAACAC 72043
QY 656 ATCCCCCACACTAGCCCTTGTCTCATCTCCAGTAAAGTTCAAGCA-CTAAGGTGGAC 714
DB 72042 A-CCCCCACACTAG-CTTTTACTCTGCTAGTAAAGATCATGATCACAACAAAGAT 71985
QY 715 ATCTCTCAATAGTATGATGCTGATATAGTTATTCCAAGCCCTCACTATACATATG 774
DB 71984 ATGATTTCTTTGTATGATATCTGATATATATATTTCAAGCCATCTTGTACTTGTG 71925
QY 775 ACTTTGAAGTGTCTACCAAGCATCTTGGTGGTTGAGAAATGGAACAGATCAGATCC 834
DB 71924 AATCTTGAAGTATCTACCAAGTATCTTGGTGGTATGAAGAACGGAACGGTCTTGATCA 71865
QY 835 AGTCATCTTACTCTCTCTTAGATATCTTGGGTTTTTGTAGGTTTTCAAATTAA 894
DB 71864 AGTACTCTCACTCTCTCTGATGATGATGATGTTTTCGTTTCTCATAAACATATA 71805
QY 895 ACATAGTCTTGTCTCTCAATGATCTCTCATATAGTCTCAATGTGTATGTTTCTCACCA 954
DB 71804 ACTCTACTTATTTCTTATATGTTTGTCTCGGATCATCTCAATGTGTATATTTCTTCA 71745
QY 955 AGGCAATGTTTGGC-----TCTTTTCATCTCTACTTCTAATATTTCTTTTGGAGCTTA 1009
DB 71744 AGGTATGATTTTGGCTTTTCTTTTCTTCTACTTCTAATAAGGATTATGTTGGAGCTCA 71695
QY 1010 GGTAGGGAATGAAGGAGGAGCATCTTGCATGTCATATGT 1050
DB 71684 GGGTAGGGGAAGAAACAAATTAATCTTATCATCATCATATCT 71644

RESULT 9
LOCUS AP003213/c
DEFINITION Oryza sativa (japonica cultivar-group) linear PLN 31-AUG-2004
BAC clone:OSJNBa0054L14.
ACCESSION AP003213 BA000010
VERSION AP003213.3 GI:14495189
KEYWORDS Oryza sativa (japonica cultivar-group)
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

```
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartioideae; Oryzeae; Oryza.
1
Sasaki,T., Matsumoto,T., Yamamoto,K., Sakata,K., Baba,T.,
Katayose,Y., Wu,J., Nilmura,Y., Cheng,Z., Nagamura,Y.,
Antonio,B.A., Kanamori,H., Hosokawa,S., Masukawa,M., Arikawa,K.,
Chiden,Y., Hayashi,M., Okamoto,M., Ando,T., Aoki,H., Arita,K.,
Hamada,M., Hayada,C., Hijishita,S., Honda,M., Ichikawa,Y.,
Idonuma,A., Iijima,M., Ikeda,M., Ikono,M., Ito,S., Ito,T., Ito,Y.,
Ito,Y., Iwabuchi,A., Kamiya,K., Karasawa,W., Katagiri,S.,
Kikuta,A., Kobayashi,N., Kono,I., Machita,K., Maehara,T.,
Mizuno,H., Mizubayashi,T., Mukai,Y., Nagasaki,H., Nakashima,M.,
Nakama,Y., Nakamichi,Y., Nakamura,M., Namiki,N., Negishi,M.,
Ohta,I., Ono,N., Saji,S., Sakai,K., Shibata,M., Shimokawa,T.,
Shomura,A., Song,J., Takazaki,Y., Terasawa,K., Tsuji,K., Waki,K.,
Yamagata,H., Yamane,H., Yoshiki,S., Yoshihara,R., Yukawa,K.,
Zhong,H., Iwama,H., Endo,T., Ito,H., Hahn,J.H., Kim,H.I., Eun,M.Y.,
Yano,M., Jiang,J. and Gojobori,T.
The genome sequence and structure of rice chromosome 1
Nature 420 (6913), 312-316 (2002)
22337376
12447438
2 (bases 1 to 150022)
Sasaki,T., Matsumoto,T. and Yamamoto,K.
Direct Submission
Submitted (19-FEB-2001) Takuji Sasaki, National Institute of
Agricultural Sciences, Rice Genome Research Program; Kamondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
On Jun 19, 2001 this sequence version replaced gi:13620982.
Genes were predicted from the integrated results of the following:
GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor
(October 1998 version). The genomic sequence was searched against
NCBI NonRedundant Protein database, nr
(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at
RGP. Protein homologs of the coding regions were searched against
NCBI NonRedundant Protein database with BLASTP2.0. ESTs represent
the identified cDNA sequences using BLASTN 2.0 with the
corresponding DBJ accession no. and RGP clone ID.
A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative-' and '-like protein'. A gene without
significant homology to any protein but with EST homology (covering
almost the entire length of partial sequence) is classified as an
'unknown' protein. A gene predicted with a gene prediction program
is classified as a 'hypothetical' protein.
The orientation of the sequence is from M13rev to -21M13 of the BAC
clone. This sequence of OSJNBa0054L14 clone has an overlap with
P0504D03 clone (DBJ: AP002970) at the position 50,801 to 150,022
of 3' end. The sequence of this clone ends at the position 99,222
of P0504D03. Detailed information on overlap and assembly quality
together with annotation of this entry is available at
http://rgp.dna.affrc.go.jp/GenomeSeq.html.
Location/Qualifiers
1..150022
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="1"
/clone="OSJNBa0054L14"
/complement(join(14480..14558,15653..15771,16436..16545,
16663..16858))
/gene="OSJNBa0054L14.1"
/complement(join(14480..14558,15653..15771,16436..16545,
16663..16858))
/gene="OSJNBa0054L14.1"
/feature="contains ESTs"
AU01313(E10180),AU161063(C53367),C19264(E10180),AU063176
(C53367),C19811(E10967)
unknown protein"
/codon_start=1
FEATURES
source
```

gene

```
/protein_id="BAB60909.1"
/db_xref="GI:14495190"
/translation="MASPAPGPKLOKAVVAVDESEFSSHALEWALRNLIAPPL
VLTQQLPLGVYASFGSPGLCTPVVAPELIKAMOQOQQLSQALLDKAKOICAHQ
VAVETWIKVQDPRKEMICQAEEKVDLLIVGSHRSPVQRLFLGVSNYCMHSHKCPV
LVVKQE"
complement(19024..19800)
/gene="OSJNBa0054L14.2"
complement(19024..19800)
/gene="OSJNBa0054L14.2"
/notes="contains ESTs"
/known protein="
AU030681(E60106), AU030680(E60106), C72766(E2205)
unknown protein"
/codon_start=1
/protein_id="BAB60910.1"
/db_xref="GI:14495191"
/translation="MSRRFLKGLTINPALASGTCOHQLOQHPVSGTAGKAKLKS
GOOLKENTIGAKKGAPSGGGGGGRGREALERTIOAESCLNASTPLRLSPKER
LEAKREELGLISKERQRELDLAKAKAKSGTRGGGVLPGLDVISGLVDED
AIPKPELTVEDGRRLAKQSVLMRRHRAQTAESLLSKKEATAALPEKLRRAAMI
PMTPPANRYATLTPPTIEGYIEKVRDAKRYSVKEKLK"
complement(123041..23226,23965..24064,24160..24356,
25077..25157,25287..25733,26015..26161)
/gene="OSJNBa0054L14.3"
complement(123041..23226,23965..24064,24160..24356,
25077..25157,25287..25733,26015..26161)
/gene="OSJNBa0054L14.3"
/notes="contains ESTs"
AU108864(C51670), C96925(C50129), C26812(C50129), C27351
(C51670), D41570(S4148)
/codon_start=1
/product="putative 26S proteasome subunit"
/protein_id="BAB60911.1"
/db_xref="GI:14495192"
/translation="MALQVVEAQORQAPDLADYADLADYQKRLWHQLKDLHFLO
LPAQGDRTIOLYNNFISDFETKLNLLKLAFAVTSQYDPKDAALFTLGGVTKL
RETRERIDPVLVYKMOJAAINLEKQDOCKLLDEKSTWSDVDPVTHASFY
WISSQYKACQFAEYKNALYLYATTVESSEFKLDLADFLDSLAALGNINIFG
ELLALPIINSLGKIVKWHMLQAFNTNLAALYDLCKVYHNAALSQAQALVONERKL
LEKINILCLMEIIFTRPESDRTPLPSVIAERTKLSIDVEYLLKMSLSVHLLEGIDE
VDSVHVSVQPVRLGIPQVKALRELDLAWGVKVTLLSVEAETPDLVAA"
join(27375..27382,27782..28478)
/gene="OSJNBa0054L14.4"
join(27375..27382,27782..28478)
/gene="OSJNBa0054L14.4"
/notes="hypothetical protein"
/codon_start=1
/protein_id="BAB60912.1"
/db_xref="GI:14495193"
/translation="MQYAGEDELGDGRSVWAVEVRRARWRGGGHVDSIGEVRTAA
HGRSOLWAEARGRRRPEAEANKRVHGDARVPGPSLSLWAGDARGHRRHSSSS
EPFRAPSHLPVPGSAMETGVGGGSTEDAGVNGDRSRTTPDPNASSVMTHGAVGGR
ACVSGSGGHDNNNACTLPVAPVPASSSSVSAGRPSPRRPRMRMLRESEAVVMESTE
QWSCDDAFSRGPIFW"
complement(28891..30563)
/gene="OSJNBa0054L14.5"
complement(28891..30563)
/gene="OSJNBa0054L14.5"
/notes="probably inactive due to frameshift(s) in CDS
probably inactive due to no initiation/termination codons
in CDS
pseudogene, gag-pol polyprotein"
pseudo
join(33259..33262,35302..35482,35726..35819)
/gene="OSJNBa0054L14.6"
join(33259..33262,35302..35482,35726..35819)
/gene="OSJNBa0054L14.6"
/notes="hypothetical protein"
/codon_start=1
/protein_id="BAB60913.1"
/db_xref="GI:14495194"
/translation="WATSGGGDARGREARRRAAVIRAGREALGRGAASSRARTGR
RARRGVTRRARQPRGLTGKAVITGGASGIGRATAEEFVPPNVF"
```

gene

CDS

gene

CDS

gene

CDS

gene

misc_feature

gene

CDS

gene

```
complement(join(37198..37348,37669..37798,37941..38118,
38273..38382,38455..38573,38690..38951,39586..39901,
40000..40296)
/gene="OSJNBa0054L14.7"
complement(join(37198..37348,37669..37798,37941..38118,
38273..38382,38455..38573,38690..38951,39586..39901,
40000..40296)
/gene="OSJNBa0054L14.7"
/notes="contains EST AU096093(S12136)
unknown protein"
/codon_start=1
/protein_id="BAB60914.1"
/db_xref="GI:14495195"
/translation="MAATSSSSPCMIASLRSHRCRLSPSATATSPPLRLTPPRRC
RNPSSSSSSQAAAIISMAPANPRHRIAPIHAAAAGGGGGGATSPSGLSLVG
IAHLVLSGIVLASDKFLKGFAPAAASIKFPSALFGVCFVSVLVLDAPALAKGFM
DFEPATLFIQRWLPFYVPSLVPLAVRDVPAASGLIKFLIIFGGWFASLMVAGYT
ALTVERKIYQOLIIPAEMKSPSPATLEFWAGAVFVAFVAVYNPATLGTATCL
PELLASTVLGMVSGIFSGVKVLUHPIICMLSADLAAIAIYGLSRSGVDVLGCDY
LTKAPNPGAGDVLGMFLGVSIIISFAFMSFKORKLVKRAAEIIFTSIAISTFSLYST
AILGRVIGLEPFLTISILPRCITVALAISIVSFEFVNSLTAAVVLTGLIGANFVQ
AAMDKLGLNDPIARGIGTASSAHGLGTAALSAKEPEALPFCIAIYGLTIGFGLICSV
PAVROSLVFIAG"
join(41004..41132,44330..44615,44748..44833)
/gene="OSJNBa0054L14.8"
join(41004..41132,44330..44615,44748..44833)
/gene="OSJNBa0054L14.8"
/notes="hypothetical protein"
/codon_start=1
/protein_id="BAB60915.1"
/db_xref="GI:14495196"
/translation="WAKTFATKTFDYVYSPNRLIPLMIGPASKYQAIPOILTREK
TKGKREKRNKTKQNRKTEKREKKTTEREPKTEKTEKQKPKQKRNKKKHARA
PPAASFACDTCFVQVGRARATGRRLAFSLKHPSPKQGLGELASLPYRTTVROGT
RTTIVR"
join(45358..45531,45640..45719,46726..46762)
/gene="OSJNBa0054L14.9"
join(45358..45531,45640..45719,46726..46762)
/gene="OSJNBa0054L14.9"
/notes="hypothetical protein"
/codon_start=1
/protein_id="BAB60916.1"
/db_xref="GI:14495197"
/translation="WTPPSATRSLLPSPRYTSLRASSLLPVVATVCPHRRGSGPF
PLEPCLDYDGGAGVDDASGYTASKSTAGILEVNRPEDESSADNHTYSYN"
```

gene

CDS

gene

CDS

Query Match

Best Local Similarity 6.7%; Score 181.4; DB 8; Length 150022;

Matches 367; Conservative 0; Mismatches 196; Indels 18; Gaps 5;

QY 486 TTTCTTGGCCCATCTTTGGTATTTTCACAAATGTCCCTCAATATGATTAATCACC

Db 123022 TTTCTTGGCATGGTGGTGGTTTGTTCATAAATTCACCTTTATACAAATAT

QY 546 ACTCATGAGCTTGTCTAGTATATAAACTCTAATCTTAAGTTTGGTGTATTTT

Db 122962 ACTTATGGAATGGTTAGTAATACACATACCTCTAAGTTGATATCTATT

QY 601 -----TGGATTTTGTGTGAAGTTGGTGGTGTAGAATAGGAGTTAAGGACC

Db 122902 TTCATTTGCTTTTATGTCATGAATTTCACATCATGATCTGGTACTTAAGG

QY 656 ATCCCCCACACTTAGCCCTTTGCTCATCTCCAGTAAAGTTCAAGGA-CTAAGGT

Db 122842 A-CCCCCACACTTAG-CTTTTACTCGTCACTGAGTAAAGATCATGATCACA

QY 715 ATCTCTCTAAATGGTACATGCTGCATATAAGTTATTTCGAAGCCTTCACCTAT

Db 122784 ATGATTCTCTTTGTAATGATCTGCATATAATTAATTTTCCAAGCCATAC

QY 775 AACTTTGAAGTGTCTACCAAGCCATCTTTGGGTGGTTGAGAAATGGAACAGAT

Db 122724 AATCTTGAAGTATCTACCAAGTCTATCTTTGGGTGGTTGAGAAACGGAAC

ACCESSION	AP004663	misc_feature	join(10557..10801,11199..11282,12931..13013,13127..13529)
VERSION	AP004663.3		/gene="P0410E02.2"
KEYWORDS	GI:38637101		/note="supported by full-length cDNA(s): AK061882"
SOURCE	Oryza sativa (japonica cultivar-group)	CDS	join(10660..10801,11199..11282,12931..13013,13127..13246)
ORGANISM	Oryza sativa (japonica cultivar-group)		/gene="P0410E02.2"
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.		/notes="contains EST(s): D2626e (C0623), C27376 (C51740), AU100814 (C51740) contains full-length cDNA(s): AK120339, AK061882, AK120022"
REFERENCE	1		/codon_start=1
AUTHORS	Sasaki, T., Matsumoto, T. and Yamamoto, K.		/product="putative ethylene-responsive transcriptional coactivator"
TITLE	Oryza sativa nipponbare (GA3) genomic DNA, chromosome 8, PAC clone: P0410E02		/protein_id="BAD03357.1"
JOURNAL	Published Only in Database (2002)		/db_xref="GI:38637102"
REFERENCE	2 (bases 1 to 144294)		/translation="MAGIPIQDPEWVVRKKAFTAAAKDEKAVNAARRSGAETET MKYNAGTNKAASGSLTKRLDDTSLAHGVSSDLKKNLMQARLDKMTQAOLA QMINKEPQVIOEYSGKAI PNQIIGKLERALGTLKRGKK"
AUTHORS	Sasaki, T., Matsumoto, T. and Yamamoto, K.	gene	15250..15480
TITLE	Submitted (23-JAN-2002) Takuji Sasaki, National Institute of Agrobiological Sciences, Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan	misc_feature	/gene="P0410E02.3"
JOURNAL	(E-mail: tsasakia@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/, Tel: 81-298-38-7441, Fax: 81-298-38-7468)		<15250..>15480
	On Dec 2, 2003 this sequence version replaced gi:31096587.	CDS	/gene="P0410E02.3"
COMMENT	Genes were predicted from the integrated results of the following: GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH (http://www.softberry.com/), GeneMark.hmm (http://opal.biology.gatech.edu/GeneMark/), GlimmerM (http://www.tigr.org/tdb/glimmerm/glmr_form.html), RiceHMM (http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4 (http://globin.cse.psu.edu/html/docs/sim4.html), gap2 (http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The genomic sequence was searched against NCBI NonRedundant Protein database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA regions were searched against NCBI NonRedundant Protein database with BLASTP. ESTs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no. and RGP clone ID. Full-length cDNAs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no.		/note="start and end point are not identified"
	A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with full-length cDNA or EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted by two or more gene prediction programs is classified as a 'hypothetical' protein according to IRGSP standard. A gene predicted by a single gene prediction program is also classified as a probable 'hypothetical' protein and is included as a miscellaneous feature of the sequence.		/note="predicted by GlimmerM etc."
	The orientation of the sequence is from T7 to SP6 of the PAC clone. This sequence of P0410E02 clone has an overlap with QJ1484_G09 (DDBJ: AP003913) clone at 5' end and with OSUNBA0091C18 (DDBJ: AP003389) at 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at http://rgp.dna.affrc.go.jp/GenomeSeq.html .		/codon_start=1
FEATURES	Location/Qualifiers		/protein_id="BAD03360.1"
Source	1..144294		/db_xref="GI:38637105"
	/organism="Oryza sativa (japonica cultivar-group)"		/translation="MRAGTAGCGFGGGGARAGGGRGARKAMTGGPHPSARAAGG PACORRARGEPMGRGERGEKWAEPQERKGGKRTSFRAMWDGDTSPRPPL PPRHATSPASLLKSSAPPDSASAPLCLLOSPTPVQRYAGSRQAPIGRL REVLPGAGSGRPHPLREERERALKRRHSAGGRSEOSIRLDLGMWEDSFINDVG EEVVVANGDGEVGMARDSDGGGNGVQDPPLSLGEVRGILEGNHILIQISQ NHKARADRLTRNALIKDLNININARVVDLYANLUTGSSHSHTSDVASATNTATNSS SPSAAAATAEVSQKPHFTTIDAK"
	/mol_type="genomic DNA"	gene	complement(27899..28693)
	/cultivar="Nipponbare"		
	/db_xref="taxon:39947"		
	/chromosome="8"		
	/clone="P0410E02"		
	complement(5113..5364)		
gene	/gene="P0410E02.1"		
misc_feature	complement(5113..5364)		
	/gene="P0410E02.1"		
	/note="hypothetical ORF predicted by GENSCAN"		
	this category is not included in IRGSP standard"		
gene	10557..13529		
	/gene="P0410E02.2"		

to base 64031. The overlap is from base 72767 to base 136858 on P0012A07. OSJNBa0034M22 also overlaps clone P0508G08 (AC130726) from base 136914 to base 149435. The overlap is from base 1 to base 12522 on P0508G08.

base 12522 on P0508G08.

Location/Qualifiers

1. .149435

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="genomic DNA"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/chromosome="5"

/clone="OSJNBa0034M22"

/clone_lib="HindIII"

538. _5565

/note="The sequence is a dinucleotide (TA) repeat in which the exact number of TA pairs is unknown. The overlapping clone P0012A07 (AC134345) shows 31 additional TA pairs compared to that which is represented by the assembly."

106220. _106269

/note="The sequence is a dinucleotide (TA) repeat in which the exact number of TA pairs is unknown. Sequences in the region show up to 13 fewer TA pairs than that which is represented by the assembly."

128275. _128340

/note="The assembly is covered by high quality sequences derived from a transposed plasmid subclone."

ORIGIN

Query Match 6.5%; Score 175.8; DB 8; Length 149435;
 Best Local Similarity 62.1%; Pred. No. 4e-32;
 Matches 378; Conservative 0; Mismatches 222; Indels 9; Gaps 6;

QY 529 ATGATAAATCCAAACATCAGGAGCTTGCTAGTTATAAATCTAAATCTTAAGTTGG 588

Db 40687 ACTCATATTTCCACCAACATAGTGAATGATTAGTAATAATCTTACCACCAAGTTGA 40746

QY 589 T-GTTTATTAGTGAGTATTGTTGTAAGTTGGTGTAGAAATAGGAGTTAAGGACCG 647

Db 40747 TACCATAATTTATTTCAATATATGCAGGTATTGATGTCATATTTATGATTTAGGACC 40806

QY 648 CCAACAGATGCCCAACATAGCCCTTTGCTCATCTCGAGTAAAGTTCAAGGACTAAG 707

Db 40807 GTCAACAACACCCCAACATTAACCTTTGCCCTCCGAGTGAAGCTATGGAATACA 40866

QY 708 GTGGAACATCTCTCAAAAGTACGATGCTGCATATAAGTTATTCAGGCTCACCTAT 767

Db 40867 ATAGAGTTTGTCTCCCTACGAAATATCATGCAACAAATATTTAAAGTCATACCTGT 40926

QY 768 ACATGT-GAACTTTGAAGTGTCTACACGCCATCTTTGGTGTGAGAAATGGAACAGAT 826

Db 40927 ACTTGTGGATCTTTGTAGTGTCTACCATGGATCCTGGCAATTGAGAAATGGAATATC 40986

QY 827 CAGATCCAGTCACTTTTACCTCTTGTGTTAGATAAATCTGGTGTTTTGAAGTTTCA 886

Db 40987 TTGACTCAAGTCACCTTAATTTCAATGAAGA-ACAACCTTGGAGTTTTCACGTAACACA 41045

QY 887 AATTTAAACATAGTCTGCTCTCAAAATGATTTCTCTATATAGTCAATGTGTATGTT 946

Db 41046 AAATTAATACTTTGCTTCTCTCTTATGATTTCTCTCG-ATCAGTCAAGGTATGATT 41104

QY 947 TCTCACAAGGCAAGTTTGTGCTCTTTTCATCTACT-TCATAATTTCTTTTGTG 1002

Db 41105 CCTCACAAGGCAATGATTGTGCTTTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 41164

QY 1003 GAGCTTAGGTAGGGAATGAAAGGAGCATCTTGCATATGCTATGTTTACTTAAGTCAA 1062

Db 41165 GAGTTAGGGAATGAAAGGAGCATCTTGTATTCATATGTTTGTATGTTGATGTTATGTT 41224

QY 1063 AACCAATCTGAGGAGCAAGTCTACATCAATCTGATC-AAGATGCGCAAGTGTGGAT 1121

Db 41225 TAATTGATCCGGAATACAGACATATCTTCTGTCGCAAAATGTCATGTGTATGTT 41284

QY 1122 ATGTGGATT 1130

Db 41285 ATGGTGAAT 41293

RESULT 13

AP005389

LOCUS

DEFINITION

Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 8, BAC clone:OSJNBa0091C18.

AP005389

ACCESSION

VERSION

AP005389.3 GI:38637257

KEYWORDS

SOURCE

ORGANISM

Oryza sativa (japonica cultivar-group)

Oryza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

1

REFERENCE

AUTHORS

Sasaki, T., Matsumoto, T. and Katayose, Y.

TITLE

Oryza sativa nipponbare (GAS) genomic DNA, chromosome 8, BAC clone:OSJNBa0091C18

JOURNAL

Published Only in Database (2002)

REFERENCE

2 (bases 1 to 154416)

AUTHORS

Sasaki, T., Matsumoto, T. and Katayose, Y.

TITLE

Direct Submission

JOURNAL

Submitted (06-JUN-2002) Takuji Sasaki, National Institute of Agrobiological Sciences, Rice Genome Research Program; Kamondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan

(E-mail: tsasaki@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/, Tel: 81-298-38-7441, Fax: 81-298-38-7468)

On Dec 2, 2003 this sequence version replaced gi:28569997.

Genes were predicted from the integrated results of the following: GENSCAN (http://ccr-081.mit.edu/GENSCAN.html), FGENESH (http://www.softberry.com/), GeneMark.hmm (http://opal.biology.gatech.edu/GeneMark/), GlimmerM (http://opal.biology.gatech.edu/GlimmerM/), RiceHMM (http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor (http://rgp.dna.affrc.go.jp/RiceHMM/), SplinePredictor (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4 (http://globin.cse.psu.edu/html/docs/sim4.html), gap2 (http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The genomic sequence was searched against NCBI Nonredundant Protein database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RGP or DBJ. Protein homologues of the coding regions were searched against NCBI Nonredundant Protein database with BLASTP. ESTs represent the identified cDNA sequences using BLASTN with the corresponding DBJ accession no. and RGP clone ID. Full-length cDNAs represent the identified cDNA sequences using BLASTN with the corresponding DBJ accession no.

A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with full-length cDNA or EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted by two or more gene prediction programs is classified as a 'hypothetical' protein according to IRGSP standard. A gene predicted by a single gene prediction program is also classified as a probable 'hypothetical' protein and is included as a miscellaneous feature of the sequence.

The orientation of the sequence is from M13rev to -21M13 of the BAC clone. This sequence of OSJNBa0091C18 clone has an overlap with P0410E02 (DBJ: AP004663) clone at 5' end and with QJ1770 H03 (DBJ: AP0052298) at 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at http://rgp.dna.affrc.go.jp/Genomeseq.html.

Location/Qualifiers

1. .154416

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="genomic DNA"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/chromosome="8"

/clone="OSJNBa0091C18"

FEATURES

source

QY 652 CAAGATCCCCACACTTAGCCCTTGTCTCATCTCGAGTAAGTTCAAGCACTAAGTGG 711
Db 5623 ACAACACCCACACTTAACTTGTCTGCGCCGAGTGAAGGCTATGACACTCTACAACAG 5682
QY 712 AACATCTCTCAAAATGTAAGTACGATCGCTGCATATAGTATTAATCCAAAGCTCACCTCATACAT 771
Db 5683 AGTTATGACTTCTTACTATGATCATGCAACGAATATCCAAAGCCATATCTGTACTT 5742
QY 772 GTGAACCTTTGAAGTGTCTACACGCCATCTTGGGTGGTGGAGAAATGAACAGATCAGAA 831
Db 5743 GTGGTCTTTGTAGTGTCTACCATGCGATCGTGAATATGAGAAATGAGAAATCTTGAC 5802
QY 832 TCCAGTCACTTTACCTCTCTTGTAGATAACTTGGGTTTTTGTAGGTTTTCAATTT 891
Db 5803 TCAAGTCACTTAATTTTCAAGAGA-ACACTTAGGGTTTTTCAAGTAACCAAAAT 5861
QY 892 AAACATAGTCTGCTCCTCAAAATGATCTCTCATATAGCTCAATGTGTATGTTCTCA 951
Db 5862 ATAACCTTGTCTCTCCTCAATGATCTCTCAATCAG-TCAAGGTGATGATCTCTCA 5920
QY 952 CCAAGGCAATGTTTGGCTCTTTCATCTCTACTCTTAATATTTCTTTTGTGGAGCTTAG 1011
Db 5921 CAAAGGCAATGATGTGCTT-----TCTTTTCTCAAGGCTTATGTGGAGCTCAG 5976
QY 1012 GTAGGCAATGAAAGCAAGCATACTTGCATTCATATGATGTTACTTAAGTCAAAACCAATC 1071
Db 5977 GTATGAAAGAGTAGAGCATTAATCTGCATCATATATCTGCAAGTCAATTAAGCATC 6036
QY 1072 TGAGGAGCAAGTATCAATCTCATCAAG-ATGTGCAAGTGTGTGG 1119
Db 6037 CATGAAATACAAGACATATCTTCTGTCAAGAATGTGCATGTGTATGG 6085

RESULT 14

OSJN00040

LOCUS

DEFINITION

Oryza sativa genomic DNA, chromosome 4, BAC clone: OSJNBb0050N09, complete sequence.

ACCESSION

AL606603

VERSION

AL606603.3

KEYWORDS

HTG.

SOURCE

Oryza sativa (japonica cultivar-group)

ORGANISM

Oryza sativa (japonica cultivar-group)

REFERENCE

AUTHORS

Feng, Q., Zhang, Y., Hao, P., Wang, S., Fu, G., Huang, Y., Li, Y., Zhu, J., Liu, Y., Hu, X., Jia, P., Zhang, Y., Zhao, Q., Ying, K., Yu, S., Tang, X., Weng, Q., Zhang, L., Lu, Y., Mu, J., Lu, Y., Zhang, L. S., Yu, Z., Fan, D., Liu, X., Lu, T., Li, C., Wu, Y., Sun, T., Lei, H., Li, T., Hu, H., Guan, J., Wu, M., Zhang, R., Zhou, B., Chen, Z., Chen, L., Jin, Z., Wang, R., Yin, H., Cai, Z., Ren, S., Lv, G., Gu, W., Zhu, G., Tu, Y., Jia, J., Zhang, Y., Chen, J., Kang, H., Chen, X., Shao, C., Sun, Y., Hu, Q., Ni, L., Zhu, F., Chen, W., Wang, L., Ding, C., Sheng, H., Gu, J., Chen, S., Li, J., Hong, G., Xue, Y. and Han, B.

TITLE

Sequence and analysis of rice chromosome 4

JOURNAL

Nature 420 (6913), 316-320 (2002)

MEDLINE

2237377

PUBMED

12447439

AUTHORS

Fu, G., Wang, S. Y., Ren, S. X., Lv, G., Lin, W., Gu, W. Q., Zhu, G. F., Tu, Y. F., Jia, J., Yin, H. F., Zhang, Y., Cai, Z., Chen, J., Kang, H., Chen, X. Y., Shao, Y., Sun, Y., Hu, Q. P., Zhang, X. L., Zhang, W., Wang, L. J., Ding, C. W., Sheng, H. H., Gu, J. L., Chen, S. T., Ni, L., Zhu, F. H., Han, S., Peng, Q., Huang, Y. C., Li, Y., Zhu, J. J., Zhao, Q., Hu, X., Liu, Y. L., Mu, J., Yu, Z., Chen, L., Fan, D. L., Weng, Q. J., Zhang, L., Lu, Y. Q., Yu, S. L., Liu, X. H., Lu, T. T., Zhang, Y. J., Lu, Y., Li, C., Li, T., Zhang, Y., Hu, H., Jia, P. X., Qian, Y. M., Ying, K., Zhou, B., Chen, Z. H., Hao, P., Zhang, L., Wu, M., Zhang, R. Q., Guan, J. P. and Hong, G. F.

DIRECT SUBMISSION

TITLE

Direct Submission

JOURNAL

Chinese Academy of Sciences, 500# Cao Bao Road, Shanghai 200233, CHINA. E-mail enquiries: bhan@ncgr.ac.cn. Clone requests: bhan@ncgr.ac.cn

REMARK

Oryza sativa japonica (nipponbare) genomic DNA, chromosome 4, BAC clone: OSJNBb0050N09.

COMMENT

On Jul 9, 2003 this sequence version replaced gi:21912461.

Web site: http://www.ncgr.ac.cn

Assembly program: phrap

Summary Statistics

This is a complete sequence.

Genes were identified by a combination of several methods: Gene prediction programs including Fgenesh (http://www.softberry.com/), genscan (http://ccr-081.mit.edu/GENSCAN.html), GeneMark-ES (http://genemark.biology.gatech.edu/GeneMark/), tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/), searches of the complete sequence against NCBI non-redundant protein database (nr) (ftp://ncbi.nlm.nih.gov/blast/db) and the ESR database at NCGR.

Location/Qualifiers

1. 127117

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="genomic DNA"

/cultivar="Nipponbare"

/sub_species="japonica"

/db_xref="taxon:39947"

/chromosome="4"

/clone="OSJNBb0050N09"

/clone_lib="CUGI-OSJNBb"

16286..17314

/gene="OSJNBb0050N09.1"

16286..17314

/gene="OSJNBb0050N09.1"

/codon_start=1

/protein_id="CAE02950.1"

/db_xref="GI:32488201"

/db_xref="UniProt/TREMBL:Q7X7R8"

/translation="MGVQEEAALAAAVKGRKQRAHAAVVPVIAASATAEESME SLSLGGAAAGSGSTTSPLLTTTAARDEAVSGVTEEDMCLMLLAGGGHGER APDAVAARAKFRSRPADGAGFEVYCKTKSCFPQALGHRTHKPKPLV APDAPEPADDKIKPTIPETAHAAAEKPKPPPPASRPATDPTVLAIPVLP QEVLDATNSAIAISVKOPRVEGSCICGAEPASGALGHMRRLPLIPASASSAVS VLDADAPROKESLSLELDLNPAPCDATAETTTTSSAATSPAFAPASDRSPLLP RALVGCYH"

complement(29123..30334)

/gene="OSJNBb0050N09.2"

complement(29123..30334)

/gene="OSJNBb0050N09.2"

/codon_start=1

/protein_id="CAE02951.2"

/db_xref="GI:38345136"

/db_xref="UniProt/TREMBL:Q7X7A0"

/translation="WRAPPPSPAVARTGFSDLQPDVGEITHGRLSFLDLRAFAVSSA ASRALKPEPLVLPGETPETATVFSGLARRSATVRASGPMRGHVIIIGFGWIVT ADERGRMLANPTVQDGLPAITIPFNATSPWGHFIMDEPPFVHIRVRGDSHW PPHYPTTAGDIRLWFRKVLVSASPRGDAAMLDDSYFGAPATAGDGRWRV APSRDVEDATHGGKFLSVTYCTGIVEAWEDGVSGETGVVTPISCPQRRKYL A RAPGRMLVLTUNGKGVKGFVQVFDENTQRWEAADIGELAILVGVNSVSVST AKHPFLTGCCVYTTDDDEIGKALRRREYGSNKPENVYSLKDGKEMSIPGLGBHL SW PPPAFPTSPFR"

complement(39012..41270)

/gene="OSJNBb0050N09.3"

complement(39012..41270)

/gene="OSJNBb0050N09.3"

/codon_start=1

/protein_id="CAE02952.1"

/db_xref="GI:32488203"

/db_xref="UniProt/TREMBL:Q7XQW6"

/translation="MAGDGMVDIIMHTLILGDMAGMGSSTELLGLSFVAQLV LTVTAGFRWRGSRMCTWFSVSDYVATTALGNLSISRTAGSRRLAFAWPFPL LHLGGPSVATAYELEDNQLSARHVLLEILLVAGAVYIVYKSTSGSWALIPASWMLFV GVAKYAEKALRRANLRVTRVERLRQRRTKANLSFAGDDDEGLMKAH TLPFKKNMVDSSVTSATDDAAIVHAKETLFRFENYKNVFRVMEMLSLMYDFLY

FEATURES
source

gene

CDS

gene

CDS

gene

CDS

TITLE

Direct Submission


```

KEYWORDS
SOURCE
ORGANISM
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
AUTHORS
Buell, C.R., Yuan, Q., Ouyang, S., Liu, J., Gansberger, K., Jones, K.M.,
Overton, II, L.L., Tsitrin, T., Kim, M.M., Bera, J.J., Jin, S.S.,
Fedorosh, D.W., Tallon, L.J., Koo, H., Ziemann, V., Hsiao, J., Blunt, S.,
Vanaken, S.S., Riedmuller, S.B., Utterback, T.T., Feldblum, T.V.,
Yang, Q.Q., Haas, B.J., Suh, B.B., Peterson, J.J., Quackenbush, J.,
White, O., Salzberg, S.L. and Fraser, C.M.
TITLE
Oryza sativa chromosome 3 BAC OSJNBa0037J17 genomic sequence
JOURNAL
Unpublished
REFERENCE
AUTHORS
Buell, R.
TITLE
Direct Submission
JOURNAL
Submitted (09-OCT-2002) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
REFERENCE
AUTHORS
Buell, R.
TITLE
Direct Submission
JOURNAL
Submitted (19-OCT-2002) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
REFERENCE
AUTHORS
Buell, R.
TITLE
Direct Submission
JOURNAL
Submitted (21-JAN-2003) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
REFERENCE
AUTHORS
Buell, R.
TITLE
Direct Submission
JOURNAL
Submitted (05-FEB-2003) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, rbuelli@tigr.org
On Jan 21, 2003 this sequence version replaced gi:24137435.
COMMENT
Address all correspondence to:rice@tigr.org
BAC clone OSJNBa0037J17 is from Oryza sativa chromosome 3
The orientation of the sequence is from Sp6 to T7 end of the BAC
clone.
Genes were identified by a combination of several methods: Gene
prediction programs including Fgenesh (http://www.softberry.com/),
Genscan and Genscan+ (Chris Burge,
http://CCR-081.mit.edu/GENSCAN.html), GeneMarkMM (Mark Borodovsky,
http://genemark.biology.gatech.edu/GeneMark/), and GeneSplicer
(Mihaela Pertea and Steven Salzberg, contact mpertea@tigr.org),
searches of the complete sequence against a peptide database and
the plant EST database at TIGR (http://www.tigr.org/tdb/tgi.shtml).
Annotated genes are named to indicate the level of evidence for
their annotation. Genes with similarity to other proteins are named
after the database hits. Genes without significant peptide
similarity but with EST similarity are named as unknown proteins.
Genes without protein or EST similarity, that are predicted by more
than two gene prediction programs over most of their length are
annotated as hypothetical proteins. Genes encoding tRNAs are
predicted by tRNAscan-SE (Sean Eddy,
http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats are
identified by RepeatMasker (Arian Smit,
http://ftp.genome.washington.edu/RM/RepeatMasker.html).
This BAC overlaps with rice BACs OSJNBa0034E08 (GB:AC135597) and
OSJNBb0047D08 (GB:AC137925).
FEATURES
source
1. .150137
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="3"
/clone="OSJNBa0037J17"
/feature="japonica cultivar-group"
230. .1537
/feature="japonica cultivar-group"
/feature="OSJNBa0037J17.1"
/feature="predicted by fgenesh"

```

```

/feature="predicted by fgenesh"
join(<230. .577,668. .1071,1210. .>1537)
/feature="OSJNBa0037J17.1"
join(230. .577,668. .1071,1210. .1537)
/feature="OSJNBa0037J17.1"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAO37479.1"
/db_xref="GI:28209461"
/translation="MEDAEPARLNDTDLVGRDRQAYVYMLSDREVANTREYSPE
LLKTKGMDFEFCAIWKAVGWOKFVVVDPSSELLTLOEFTLKEIEDGIESPPEHKEY
TLTWKGLSTLGFHFKPRNDINPTLRILHKHMIANTWPRGDLRPIREDELLIMFAMV
RKILVPMKIRQWLESKFSAPVECTNEILIRIAGKGLVSDQAFATISATPCIDET
YLVQGHIVKHGINGSLLIYFFPGCTNEIFLPNAGCMRQVSGQLGHPDGSKLRATASES
QAGRVLRTGGVRLMTYIRGTNTLPDQVTEIYQNLMEHIAQTOEWQSQSVNAQFASIN
NNMQQHDLDLQAYFRFGQFNPYQGP"
1576. .1621
/rpt_family="tandem"
/complement(4405. .5112)
/feature="OSJNBa0037J17.2"
/feature="similar to gypsy type retrotransposon protein orf3
GB:BAA84461 GI:5902448 (Oryza sativa (japonica cultivar
group))"
/complement(<4405. .>5112)
/feature="OSJNBa0037J17.2"
/complement(4405. .5112)
/feature="OSJNBa0037J17.2"
/codon_start=1
/product="putative retrotransposon protein"
/protein_id="RAO37480.1"
/db_xref="GI:28209462"
/translation="MAPCKPTNAGATGDDPGSLGEGWGCYLGKSLVKESDLADLILTS
VLAEGQATCWAKAVVSPGDRRTVFAAFFTGLRLPCDDFLPSVLEKLPQLSP
LAFPKLAIFAWCRCTGLAPTAELFAVLFTACTTTQVNTVHAGRAEEDSFLRLIYAS
AGAFRLAGACVDGKMGQPLDAEVLVRQRPILHRQGLAPLSAPSCFHRQAKCRD
RHRTSPGSGSFPTLRGG"
7834. .8349
/feature="OSJNBa0037J17.3"
/feature="predicted by GenemarkHMM"
join(<7834. .7915,8249. .>8349)
/feature="OSJNBa0037J17.3"
join(7834. .7915,8249. .8349)
/feature="OSJNBa0037J17.3"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAO37477.1"
/db_xref="GI:28209459"
/translation="MAKTIKQWRRLKDRAGIQDQMLGQIVASTAGAPGDLRAK
ALRNATKGRKGGKC"
complement(7844. .10311)
/feature="OSJNBa0037J17.4"
/feature="predicted by fgenesh"
complement(join(<7844. .8124,9092. .>9316,9474. .>9567,
9622. .9897,9973. .>10311))
/feature="OSJNBa0037J17.4"
complement(join(7844. .8124,9092. .9316,9474. .>9567,
9622. .9897,9973. .>10311))
/feature="OSJNBa0037J17.4"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAO37481.1"
/db_xref="GI:28209463"
/translation="MAPCKNPFASVMPGDDPRIDDDTTAYLGASIVDEDAKLVTSQ
VLAEQAFTPGKAVVPLKGDNRVTFTVFEEAGRLFCNVLPLRILRHVELPOLSP
SALFRMAIFDWTVVTVVTKMKKMAFGSMFNVPERSDLVPVNTKMSWDCWMSKWFY
HSIIPFEASAKALRWRRAIALNRKPKVTDGMEARFTLLRKIPFLFTQVFEVS
PNEEDVGLPKVLMLEGANASCLIRGCPAGCTWANFLPRAKPKFADGDTGNGN
SGKNLPLGSGPLPPPTMTLTKTLCMTWTKRKEVPPLGFRILRYEITMAALMKGTG
AWVYSVIWIGHPLANGQLANRKPQLGWLPGPLQSRACYGEAIPLOTHPAQDAFLARY
ORDPSAFAISAL"
10940. .11464
/feature="OSJNBa0037J17.5"
/feature="predicted by fgenesh"

```



```
mrna
  join(<10940..10998,11443..>.11464)
  /gene="OSJNBa0037J17.5"
  join(10940..10998,11443..11464)
  /gene="OSJNBa0037J17.5"
  /codon_start=1
  /product="hypothetical protein"
  /protein_id="AAO37482.1"
  /db_xref="GI:28209464"
  /translation="MAKAEGSRWYLARPNASWAKSVVPMQ"
repeat_region
  12527..12583
  /rpt_family="CT-rich"
  /complement(12586..14241)
  /gene="OSJNBa0037J17.6"
  /note="predicted by genemarkHMM"
  complement(join(<12586..12666,12994..13344,13455..13628,
13822..13878,13954..>.14241))
  /gene="OSJNBa0037J17.6"
  complement(join(12586..12666,12994..13344,13455..13628,
13822..13878,13954..14241))
  /gene="OSJNBa0037J17.6"
  /codon_start=1
  /product="hypothetical protein"
  /protein_id="AAO37494.1"
  /db_xref="GI:28209476"
  /translation="MATVTESWARLAASGGAGDGRRRWAWPAAAEAGGRWWE
VHDDGLKBAKVAATLMMWRDEARTPTVRLTGKRESKLSSEALAAALDPGGKVALGWA
SSGGRWGRGGRPQRPWIPAAERSAIDGLLASRHGVDRVTVDGTARSRSGRRAPEL
APEFDGSLRWKVKAAERSLWHDLASDSVDSIDSAYADSSTSADSASTGD
SASTMPGEVKRAARVPGRFRRHGLTAHETGRAVTHVPSRRPRKGPW
MAGPCRHGMRVARRAVPGHAGYFRFRASGRPFPPFGHL"
  12719..12746
  /rpt_family="GC-rich"
  repeat_region
    13100..13137
    /rpt_family="tandem"
  repeat_region
    13111..13169
    /rpt_family="tandem"
  repeat_region
    13796..13852
    /rpt_family="tandem"
  repeat_region
    14242..14283
    /rpt_family="G-rich"
  repeat_region
    14349..14420
    /rpt_family="tandem"
  14705..18463
  /gene="OSJNBa0037J17.7"
  /note="highly similar to ribosomal protein S5 GB:CAA63650
GI:5725342 (Spinacia oleracea)"
  join(<14705..15435,18214..>.18463)
  /gene="OSJNBa0037J17.7"
  join(14705..15435,18214..18463)
  Query Match          6.4%; Score 172; DB 8; Length 150137;
  Best Local Similarity 59.0%; Pred. No. 3.6e-31;
  Matches 385; Conservative 0; Mismatches 260; Indels 7; Gaps 5;
  QY 485 TTTTCTGGCCCATCTTTGGTATTTTCACAAATGCCCCCTACAAATGATAATCACCAC 544
  DB 30628 TTTTCTGCATGACAGATGTTTTTCTCTCAATTAACCTACATACATCTTTCCACCA 30687
  QY 545 AACTCATGAGCTTGGTAGTATAAATCTTAATCTTAAGTTGGT-GTTTATTGAGTGG 603
  DB 30688 CACTAGTAGAAATATTAGTAATAATCCCTACCACTAAGTTGGATACCATATTTTATTGC 30747
  QY 604 ATTTTGTGTGAAGTTGGTGTAGAAATAGGATTAAGGACCGCCCAACAGATCCCCCA 663
  DB 30748 ATTATATGAGGTATTACGGTCAATATATTGATTTAAGGACCGTTAGCAACACCCCA 30807
  QY 664 CACTTAGCCCTTTGCTCATCTCGAGTAAAGTTCAAGGACTAAGGTGGAACATCTCTCTCA 723
  DB 30808 CACTTAACCTTTGCTCATCCCGAGTGAAGGCTATGGACTCTACACAAAGTTACGACTTC 30867
  QY 724 AATGTACGATCGCTGCATATAAGTTATTCCAGGCTCACCTATACATGT-GAACTTTGA 782
  DB 30868 CTTACTATAAATCAGCTAAACAAATATTTCAAAAGTCATACCTATATCTTGTGGATCTTTGT 30927
  783 AGTGCTTACACGCCCATCTTTGGGTGGTTGAGAAATGGAAACAGATCAGAAATCCAGTTCATCT 842
  30928 AGTGCTTACCATGCAATCTTGAGCAATTGAGAAATGGAAACATCTTGACTTAACTAACCC 30987
  843 TTACCTCTCTTGGCTTAGATAACTTGGGTTTTTGAAGTTTTTCAAAATTTAAACATAGTC 902
  30988 TAACTTTTTCATGAAGAACAACTTGGGGTTTTTCAAGTAAACACAAATTTATAACTTTGCT 31047
  903 TTGCTCTCTCAATGATTTCTCTCATATAGCTCAATGTGTGTGTTTCTCACCAGGCAATG 962
  31048 TTACTCTCTCGATTGATTCTCTC-GATTACTCAAGTGTATGATTTCTCAAGGCAATAT 31106
  963 TTTTGCCTCTTTTCAT---CCTACTTCTAATATTTTCTTTTGGAGCTTAGGTTAGGGAA 1019
  31107 GATTGTGCTTTTCTTTTCTCTACCATCTAATAAGGTTTATGTGGAACCTTAGGGTATAGAA 31166
  1020 TGAAGAAGGAGATACCTTGGCATGTATGTTACTTAAGTCAAAACCAAAATCTGAGGAGA 1079
  31167 AAGAGTAGAGCATTTGCTTGCATCATATATTGCAAGTCAAAATATGATCCACGAAA 31226
  1080 AGCAAGTCATACAAATCTGTATCAAG-ATGTGCAAGTGTGTGGATATGTGGATT 1130
  31227 TACAAGACATACCTTTCTGGTCAAGAAATGTGAATGTGTATGTGTACGGTGAAT 31278
  Search completed: May 9, 2005, 01:51:44
  Job time : 11595 secs
```

THIS PAGE LEFT BLANK

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2005, 21:22:28 ; Search time 1413 Seconds
(without alignments)

11269.713 Million cell updates/sec

Title: US-10-751-550-1

Perfect score: 2690

Sequence: 1 tctagataatagactact.....ccataactagtagacaccatgg 2690

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001s:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2690	100.0	2690	13	ADR21934 Sugarcane
2	1001.2	37.2	1047	4	AAS01032 Sugarcane
3	151.4	5.6	2000	12	ADJ41549 Plant cDN
4	149.2	5.5	4797	12	ADJ40495 Plant cDN
5	114.4	4.3	2000	12	ADJ41110 Plant cDN
6	55.4	2.1	2000	8	ADA71938 Rice gene
7	53.4	2.0	51259	2	AAX83007 Partial m
8	52.2	1.9	72352	12	ADQ97067 Mouse can
9	51.8	1.9	3718	2	AAV72244 G. max SB
10	51	1.9	1150	8	ABZ57867 Porcine e
11	51	1.9	2860	8	ABZ57861 Porcine e
12	50.8	1.9	1976	8	ABZ57865 Porcine e
13	50	1.9	267	4	AAK45604 Human bon
14	50	1.9	267	4	AAK19599 Human bra
15	50	1.9	267	4	ABSA45294 Human liv
16	50	1.9	267	6	ABS19876 Human gen
17	50	1.9	474	4	AAK32380 Human bon
18	50	1.9	474	4	AAK06682 Human bra
19	50	1.9	474	4	ABS32088 Human liv
20	50	1.9	474	6	ABS07163 Human gen

21	50	1.9	943	12	ADM72367
22	50	1.9	2385	3	AAC83331
23	50	1.9	3809	6	ABK95303 Human pr
24	49.6	1.8	2623	4	AAD17484 Mouse glu
25	49.6	1.8	3815	8	ABS57008 Genomic D
26	49.2	1.8	795	2	AAV55830 FLGA inse
27	49.2	1.8	799	2	AAV55831 Nucleotid
28	49.2	1.8	1925	2	AAx90924 Epstein B
29	49.2	1.8	1926	3	AAAS0254 Epstein B
30	49.2	1.8	1926	4	AAf82902 EBV tethe
31	49.2	1.8	1926	10	ADK65580 Human her
32	49.2	1.8	2580	3	AAa75454 Nucleotid
33	49.2	1.8	2580	6	AAI64275 Epstein-B
34	49.2	1.8	5452	2	AAx90923 Anti-sens
35	49.2	1.8	8705	2	AAz23778 Vector ps
36	49.2	1.8	8705	12	ADM10659 Expressio
37	49.2	1.8	9482	12	ADP64415 Vector pc
38	49.2	1.8	9600	2	AAV21683 Vector pl
39	49.2	1.8	10285	6	ABS71027 pCEP-Xa-F
40	49.2	1.8	10285	6	ABS66453 Plasmid p
41	49.2	1.8	10330	12	ADL67154 Plasmid p
42	49.2	1.8	10380	2	AAZ22248 Nucleotid
43	49.2	1.8	10477	12	ADL67152 Plasmid p
44	49.2	1.8	10516	12	ADL67150 Plasmid p
45	49.2	1.8	10561	12	ADL67148 Plasmid p

ALIGNMENTS

RESULT 1

ADR21934

ID ADR21934 standard; DNA; 2690 BP.

XX ADR21934;

XX 21-OCT-2004 (first entry)

DE Sugarcane jasmonate-induced protein (JAS) promoter DNA Seqid 1.
XX jasmonate-induced protein; JAS; promoter; stem-specific;
KW defence-inducible; carbon metabolism; insecticidal; pest tolerance;
KW plant; ds.

XX Saccharum.

XX Key

FT CAAT_signal

FT Location/Qualifiers

FT 2428..2431

FT /*tag= a

FT TATA_signal

FT 2604..2610

FT /*tag= b

FT misc_feature

FT 2687..2689

FT /*tag= c

FT /note= "Start codon (AUG) of JAS"

FT WO2004062366-A2.

XX 29-JUL-2004.

XX 05-JAN-2004; 2004WO-US000115.

XX 03-JAN-2003; 2003US-0437974P.

XX (TEXA) UNIV TEXAS A & M SYSTEM.

XX Damaj MB, Mirkov ET, Reddy AS, Thomas TL, Rathore KS, Emani C;

XX Kumpatla SP;

XX WPI; 2004-544019/52.

XX New isolated nucleic acid comprises jasmonate-induced protein (JAS)
PT promoter and an exogenous nucleic acid, useful as a promoter for altering
PT carbon metabolism in a plant cell or for driving expression of

PT insecticidal proteins in sugarcane.

XX Claim 1; SEQ ID NO 1; 58pp; English.

XX This invention relates to a novel nucleic acid molecule that comprises a
CC jasmonate-induced protein (JAS) promoter and an exogenous nucleic acid.
CC Specifically, it refers to a stem-specific promoter active in plants that
CC exhibits enhanced specificity in regulating gene expression in stem
CC tissues and in response to induction by external stimuli such as plant
CC defence-inducing agents. The present invention describes a bacterial cell
CC comprising an expression vector that can be used to transform a monocot
CC plant such as sugarcane, sorghum, rice, maize or hybrids thereof, in
CC order to upregulate localised expression. In particular, this nucleic
CC acid is useful as a promoter for altering carbon metabolism in the
CC sucrose accumulating tissues, and for driving expression of insecticidal
CC proteins in sugarcane. These promoters may also be applied to the
CC development of improved pest and disease tolerant rice plants. This
CC polynucleotide sequence is the JAS promoter DNA sequence of the
CC invention.

XX
SQ Sequence 2690 BP; 732 A; 608 C; 586 G; 764 T; 0 U; 0 Other;

Query Match 100.0%; Score 2690; DB 13; Length 2690;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2690; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTAGATAATACGACTCATTATAGGCGACGCGTGTGCGACCGGCCGGGTGTCGGAC 60

DB 1 TCTAGATAATACGACTCATTATAGGCGACGCGTGTGCGACCGGCCGGGTGTCGGAC 60

QY 61 AGCTAGAGGGCGCCACCGCGCTCTAGCTTCTCCAACTTCTCGTCGAGATCCCTTCAGGG 120

DB 61 AGCTAGAGGGCGCCACCGCGCTCTAGCTTCTCCAACTTCTCGTCGAGATCCCTTCAGGG 120

QY 121 ATGCCCAATGCCACCGCCCTTAAGTCAACTGCGGAGCTGGAGCTTCGCCAGGGTCAGA 180

DB 121 ATGCCCAATGCCACCGCCCTTAAGTCAACTGCGGAGCTGGAGCTTCGCCAGGGTCAGA 180

QY 181 GCTGCGGCGACCGCTGGTAGCCGATTCCTGATGACCGCGGGGTGCGTCCATGAAG 240

DB 181 GCTGCGGCGACCGCTGGTAGCCGATTCCTGATGACCGCGGGGTGCGTCCATGAAG 240

QY 241 AAGTCATTTCGCCCAACCAAGTCGAGTGGTCCCTGAGGGGGGGGGAAGCAAAAGCT 300

DB 241 AAGTCATTTCGCCCAACCAAGTCGAGTGGTCCCTGAGGGGGGGGGAAGCAAAAGCT 300

QY 301 TGCATGCACTAGCGCCCTGGCAGCGAGCTCCTGTAGTATCACTCGCTCGCTCCAGCT 360

DB 301 TGCATGCACTAGCGCCCTGGCAGCGAGCTCCTGTAGTATCACTCGCTCGCTCCAGCT 360

QY 361 CATGCTCGCAAGCTCCAGGGGGCGCGAGTGTCTCCAACTTCGACCTTCCTCTTCA 420

DB 361 CATGCTCGCAAGCTCCAGGGGGCGCGAGTGTCTCCAACTTCGACCTTCCTCTTCA 420

QY 421 GCTCTCTTCCACATGCACTGCTGCTCGCAGCGACCTTCTCCACCTTTTACTCTTTTCT 480

DB 421 GCTCTCTTCCACATGCACTGCTGCTCGCAGCGACCTTCTCCACCTTTTACTCTTTTCT 480

QY 481 TCTCTTTTCTTGCGCCCATCTTTGGTATTTTCAAAATGTCGCCCTTCAAAATGATAATCA 540

DB 481 TCTCTTTTCTTGCGCCCATCTTTGGTATTTTCAAAATGTCGCCCTTCAAAATGATAATCA 540

QY 541 CCAAAACTCATGAGCTCTAGTTATATAAATCTAATCTTAAGTTTGGTGTATTTTGGAG 600

DB 541 CCAAAACTCATGAGCTCTAGTTATATAAATCTAATCTTAAGTTTGGTGTATTTTGGAG 600

QY 601 TGGATTTTGTGTAAGTTGGTGGTTAGAAATAGGATTTAAGGACCGCCCAACAAGATCCC 660

DB 601 TGGATTTTGTGTAAGTTGGTGGTTAGAAATAGGATTTAAGGACCGCCCAACAAGATCCC 660

QY 661 CCACACTTAGCCCTTGTCTCATCTCGAGTAAAGTTCAAGGACTAAGGTGGAAACATCTCC 720

DB 661 CCACACTTAGCCCTTGTCTCATCTCGAGTAAAGTTCAAGGACTAAGGTGGAAACATCTCC 720

QY 721 TCAAAATGGTACGATCGCTGCATATAGTTATTTCCAAAGCCTCACCTATACATGTGAATTT 780

DB 721 TCAAAATGGTACGATCGCTGCATATAGTTATTTCCAAAGCCTCACCTATACATGTGAATTT 780

QY 781 GAAGTGTCTACACCGCATCTTGGGTGGTTGAGAAATGGAACAGATCAGAAATCCAGTCAAT 840

DB 781 GAAGTGTCTACACCGCATCTTGGGTGGTTGAGAAATGGAACAGATCAGAAATCCAGTCAAT 840

QY 841 CTTTACCTCTCTTGTCTAGATAAATCTTGGTGGTTTCTTAAGGTTTCAAAATTTAAACATAG 900

DB 841 CTTTACCTCTCTTGTCTAGATAAATCTTGGTGGTTTCTTAAGGTTTCAAAATTTAAACATAG 900

QY 901 TCTTGTCTCTCAAAATGATTTCTCTCATATAGTCAATGTGTATGGTTTCTCACCAAGGCAA 960

DB 901 TCTTGTCTCTCAAAATGATTTCTCTCATATAGTCAATGTGTATGGTTTCTCACCAAGGCAA 960

QY 961 TGTGTTGGCTCTTTTTCATCCTTCTTAATATTTCTTTTGTGGAGCTTAGGGTAGGGAT 1020

DB 961 TGTGTTGGCTCTTTTTCATCCTTCTTAATATTTCTTTTGTGGAGCTTAGGGTAGGGAT 1020

QY 1021 GAAAAGGAAGCATACTTGCATTGCATATGTTTACTAAGTCAAAACCAAAATCTGAGGAGAA 1080

DB 1021 GAAAAGGAAGCATACTTGCATTGCATATGTTTACTAAGTCAAAACCAAAATCTGAGGAGAA 1080

QY 1081 GCAAGTCATCAATCTGATCAAGATGTGCAAGTGTGCGATATGTGGATTAAGATAACTC 1140

DB 1081 GCAAGTCATCAATCTGATCAAGATGTGCAAGTGTGCGATATGTGGATTAAGATAACTC 1140

QY 1141 CTGTTTATTCATGCTCTCTCTCTTAATAAATCTTTAGAGGGCATGCAATCTTTGATGGG 1200

DB 1141 CTGTTTATTCATGCTCTCTCTCTTAATAAATCTTTAGAGGGCATGCAATCTTTGATGGG 1200

QY 1201 CTTTCATGAGCTCATCTGATGTCTAAGCATGAGCTCATCAATTTTATAAGCATGGTCAAT 1260

DB 1201 CTTTCATGAGCTCATCTGATGTCTAAGCATGAGCTCATCAATTTTATAAGCATGGTCAAT 1260

QY 1261 ACCAAATTTACTCTTTTGGAGCATGTTTATTTAGGAGGAGCTTTTACCTGTTGAGGTA 1320

DB 1261 ACCAAATTTACTCTTTTGGAGCATGTTTATTTAGGAGGAGCTTTTACCTGTTGAGGTA 1320

QY 1321 AATCTGAAAGCTTAATAAATCGGCTAAGCAAAATATTTTACCTGTTGATTTCTAACCAAT 1380

DB 1321 AATCTGAAAGCTTAATAAATCGGCTAAGCAAAATATTTTACCTGTTGATTTCTAACCAAT 1380

QY 1381 TTGATGTAGGCAATATTTGATGAGTGACTGCAAAATGATTTGAAGGCTTTTAAAGGAGATT 1440

DB 1381 TTGATGTAGGCAATATTTGATGAGTGACTGCAAAATGATTTGAAGGCTTTTAAAGGAGATT 1440

QY 1441 GAGAGGATTAATCTACAAATTAATAATGTAAGAGAAAGCATTTCAAGTGTGAGATCTGG 1500

DB 1441 GAGAGGATTAATCTACAAATTAATAATGTAAGAGAAAGCATTTCAAGTGTGAGATCTGG 1500

QY 1501 TGTGGAAGACTATTTTGGCTCTTGGGGGTAAAGACAAAGCTTTAGTAAAGTGGCTCAA 1560

DB 1501 TGTGGAAGACTATTTTGGCTCTTGGGGGTAAAGACAAAGCTTTAGTAAAGTGGCTCAA 1560

QY 1561 AATTGGGAGGGCCCATGCAAGATTTGATTAAGTAATTTTGGATTTGAGGAGGCAATTTTC 1620

DB 1561 AATTGGGAGGGCCCATGCAAGATTTGATTAAGTAATTTTGGATTTGAGGAGGCAATTTTC 1620

QY 1621 AAGGTGATCATCTACCTAGAGCTCTCAATGGAGGTGCTCGAAGACATATTACCCATGTG 1680

DB 1621 AAGGTGATCATCTACCTAGAGCTCTCAATGGAGGTGCTCGAAGACATATTACCCATGTG 1680

QY 1681 TATGCAAGAGTGTGTAGCTAGTAACTGATGTGTAACGATCTCCAATGGGGCAAG 1740

DB 1681 TATGCAAGAGTGTGTAGCTAGTAACTGATGTGTAACGATCTCCAATGGGGCAAG 1740

QY 1741 ACATATTACCTTAAGCCAGGCTGTTTGGAGTTCCAGTTCCAGTAGGATATAGAGATTCTCGT 1800

DB 1741 ACATATTACCTTAAGCCAGGCTGTTTGGAGTTCCAGTTCCAGTAGGATATAGAGATTCTCGT 1800

Db 301 CCAACCTAATTAATAGTTCCTTCCTTGGCTCTAGGACAAATGACGTGTTCCGGTATCC 360
QY 1999 TGTGTAATTAAGAACAAACCTAGTGTGACCTGTCCGATAGAGTCCACCTGGGTAGGCA 2058
Db 361 TGTCTGAATCAAGAACAAACCTAGTGTGACCTGTCCGATAGAGTCCACCTGGGTAGGCA 420
QY 2059 TTTATAGGATTCGTGTATTTCTCTGCAAAAAGCGATTAGCTGGCTTTAAACCTGGCT 2118
Db 421 TTTATAGGATTCGGGTATTTCTCTGCAAAAAGCGATTAGCTGGCTTTAAACCTGGCT 480
QY 2119 AGCCCGGATTCGTGGCTTCCTACCTACCAAGTGAATTTTCATGTGATCGGTGATTTAGCA 2178
Db 481 AGCCCGGATTCGTGGCTTCCTACCTACCAAGTGAATTTTCATGTGATCGGTGATTTAGCA 540
QY 2179 CTTTGTATGTAACCCAACTTAAGTCGACAACTATAAATATGCTACTTTCAGAGATGTTA 2238
Db 541 CTTTGTATGTAACCCAACTTAAGTCGACAACTATAAATATGCTACTTTCAGAGATGTTA 600
QY 2239 TCACGACACAACTCCTAATCTACGGAAGCTTAAGTTAGTTTGTCTCGGAGACAAGCAAT 2298
Db 601 TCACGACACAACTCCTAATCTACGGAAGCTTAAGTTAGTTTGTCTCGGAGACAAGCAAT 659
QY 2299 TGTGGCCAGTCACTATAGCTTACCTCAGAGGAGTGTGGAGCAGTTCGCTGTTGGATTGA 2358
Db 660 TGTGGCCAGTCACTATAGCTTACCTCAGAGGAGTGTGGAGCAGTTCGCTGTTGGATTGA 719
QY 2359 AAACAGGTGATCGTATCAGATTAATGATTCACATGACAGTAATAATGTGGTACAGTAA 2418
Db 720 AAACAGGTGATCGTATCAGATTAATGATTCACATGACAGTAATAATGTGGTACAGT-A 778
QY 2419 CTTTCGCAACAAATAAATCTGTACAAATTTATAGTGCATCTCTGACGTAAATGCTTC 2478
Db 779 CTTTCGCAACAAATAAATCTGTACAAATTTATAGTGCATCTCTGACGTAAATGCTTC 838
QY 2479 TAGCTCAGAGGATTTGATTCAGAGGCGCTGACCCATCACTAATGACGGCTTTTACCC 2538
Db 839 TAGCTCAGAGGATTTGATTCAGAGGCGCTGACCCATCACTAATGACGGCTTTTACCC 898
QY 2539 ATCATATGACCAATTTGTCATATCATGATGATTCATGTCGCTCTGTCATGACGACGAG 2598
Db 899 ATCATATGACCAATTTGTCATATCATGATGATTCATGTCGCTCTGTCATGACGACGAG 958
QY 2599 CCCTCTATAAATACTGGCATCTCCCTCCCGTTCCACAGATCACACACACACAGCAAT 2658
Db 959 CCCTCTATAAATACTGGCATCTCCCTCCCGTTCCACAGATCACACACACACAGCAAT 1018
QY 2659 AAACGGTAGTGCCTAATAGTAGTACA 2684
Db 1019 AAACGGTAGTGCCTAATAGTAGTACA 1044

RESULT 3

ADJ41549
ID ADJ41549 standard; cDNA; 2000 BP.

XX AC ADJ41549;

XX DT 06-MAY-2004 (first entry)

XX DE Plant cDNA #2549.

XX KW Plant; gene; ss; transcription; plant genome augmentation; cereal;
KW soybean; alfalfa; sunflower; canola; cotton; peanut; tobacco; sugar beet;
KW maize; barley; sorghum; rice; wheat; crop plant; insecticide resistance;
KW stress tolerance; salt tolerance; cold tolerance; drought tolerance;
KW plant nutrition; apical dominance; dwarfism; early flowering; antiviral;
antifungal.

XX OS Eukaryota.

XX PN US2004016025-A1.

XX XX

PD XX 22-JAN-2004.
PF XX 26-SEP-2002; 2002US-00260238.
XX 26-SEP-2001; 2001US-0325277P.
PR 26-SEP-2001; 2001US-0325448P.
PR 04-APR-2002; 2002US-0370620P.
XX (BUDW/) BUDWORTH P.
PA (MOUG/) MOUGHAMER T.
PA (BRIG/) BRIGGS S P.
PA (COOP/) COOPER B.
PA (GLAZ/) GLAZERBROOK J.
PA (GOFF/) GOFF S A.
PA (KATA/) KATAGIRI F.
PA (KREP/) KREPS J.
PA (PROV/) PROVANT N.
PA (RICK/) RICHE D.
PA (ZHUT/) ZHU T.
XX Budworth P, Moughamer T, Briggs SP, Cooper B, Glazebrook J;
PI Goff SA, Katagiri F, Kreps J, Provant N, Ricke D, Zhu T;
XX WPI; 2004-190374/18.
XX New rice promoter, useful for manipulating crop plants to alter or
PT improve phenotypic characteristics, e.g. produce large quantities of oil
PT or proteins, resistance to insecticides, virus or fungi, stress tolerance
PT or high nutritional value.
XX Claim 1; SEQ ID NO 2549; 230pp; English.
XX The invention relates to plant nucleotide sequences that direct seed-,
CC leaf- and/or stem-, panicle-, root- or pollen-specific or -preferential
CC or constitutive transcription of an operatively linked nucleic acid
CC segment. The invention also relates to a method for augmenting a plant
CC genome and a method of identifying a gene, where its expression is
CC altered in the seed, leaf, stem, panicle, pollen, root or is constitutive
CC in a plant cell. The plant is a cereal, e.g. soybean, alfalfa, sunflower,
CC canola, cotton, peanut, tobacco or sugar beet, preferably maize, barley,
CC sorghum, rice or wheat. The polynucleotides and the polypeptides they
CC encode are useful for manipulating crop plants to alter or improve
CC phenotypic characteristics, to produce large quantities of oil or
CC proteins, to incur resistance to insecticides, viruses or fungi, and to
CC incur stress tolerance (e.g. salt, cold or drought) to ensure the plants
CC have a high nutritional value with reduced apical dominance or dwarfism,
CC early flowering or altered metabolic pathways. This sequence represents a
CC plant nucleic acid of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification but was obtained in
CC electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.
XX SQ Sequence 2000 BP; 537 A; 374 C; 551 G; 536 T; 0 U; 2 Other;
Query Match 5.6%; Score 151.4; DB 12; Length 2000;
Best Local Similarity 61.6%; Pred. No. 3.3e-30;
Matches 305; Conservative 0; Mismatches 166; Indels 24; Gaps 3;
QY 1232 GAGCTCATCTTTATATAGCATGTGTATACCAAAATTTACTCTTTTGAGCATGTTTATA 1291
Db 86 GGGCGCATAGGATATCTAAACATGTGTGCCACTTAAAGTTTACTCTTTTGAGTTGTTATG 145
QY 1292 TTTAGGAGGAGCTTTTACTCTGTGTAGGTAAATCTCAACGCTAAATAATCGGCTAAGCAA 1351
Db 146 GTCAAGAGAGCCGTTTTTACGGTGGAGGAAATCTTGGTCTCTCTGTTATATCAAGCAAG 205
QY 1352 ATAAATTTATCACTCTGTTGATCTTAACAATTTGATGACGACAAATTTGATGAGGTGACTG 1411
Db 206 ATGATTTGTCAGTGAAGATTACAGATGTTTGTGGGAGACAAATCTTGTATGAAGTCATCG 265
QY 1412 ACAATGATTTGAGGCTTTAAGGAGATTTGAGAGATTAAGAGAGAGAGAGAGAGAGAGAG 1455
Db 266 ACAAGCGCTTGAAGGCAATTTGGAAGAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 325


```
XX WO2003002746-A2.
XX
XX
PD 09-JAN-2003.
XX
XX
PF 28-JUN-2002; 2002WO-EP007159.
XX
XX
PR 29-JUN-2001; 2001US-0302133P.
XX
XX
PR (NOVS ) NOVARTIS AG.
XX
PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.
XX
XX
PI Herring CT, Langford G, Quinn G, Scobie L;
XX
XX
DR WPI; 2003-210279/20.
XX
XX
PT An isolated polypeptide useful for screening pigs for the presence of a
PT porcine endogenous retrovirus (PERV) and providing a pig modified not to
PT express a PERV.
XX
XX
PS Claim 1; Page 45-47; 63pp; English.
XX
XX
CC The present sequence is that of a claimed flanking sequence of porcine
CC endogenous retrovirus (PERV). Flanking sequences are DNA sequences
CC surrounding a PERV which represent a unique molecular signature that can
CC be used to characterise PERV integration sites and identify PERVs. The
CC present sequence is one of eight claimed PERV polynucleotides (see
CC ABZ57861-68) useful for screening a pig for the presence of PERV. The
CC screening provides for the elimination of donors with known proviruses. A
CC pig modified not to express a selected PERV is obtained by identifying a
CC PERV using the claimed polynucleotides, and knock-out or inactivation of
CC the PERV by homologous recombination. The organs, tissues and cells of
CC the modified pig are suitable for use in xerotransplantation
XX
SQ Sequence 1150 BP; 333 A; 265 C; 247 G; 305 T; 0 U; 0 Other;

Query Match 1.9%; Score 51; DB 8; Length 1150;
Best Local Similarity 91.5%; Pred. No. 0.0075;
Matches 54; Conservative. 0; Mismatches 5; Indels 0; Gaps 0;

QY 7 TAATACGACTCTACTATAGGCGACGCGTGTGTCGACGCGCGGCTGTCTGCGACAGCTA 65
DB 54 TAATACGACTCTACTATAGGCGACGCGTGTGTCGACGCGCGGCTGTCTGCGACCA 112

RESULT 11
ABZ57861
-ID ABZ57861 standard; DNA; 2860 BP.
XX
XX
AC ABZ57861;
XX
XX
DT 14-APR-2003 (first entry)
XX
DE Porcine endogenous retrovirus flanking sequence.
XX
KW PERV; knockout animal; pig; xerotransplantation; organ transplant; ss.
XX
OS Porcine endogenous retrovirus.
XX
FH Key Location/Qualifiers
FT misc_feature 1..21
FT /tag= a
FT /note= "vector"
FT primer_bind 2536..2556
FT /tag= b
FT LTR 2661..2860
FT /tag= d
FT primer_bind 2661..2678
FT /tag= c
FT primer_bind 2839..2858
FT /tag= e
XX
PN WO2003002746-A2.

XX
XX
PD 09-JAN-2003.
XX
XX
PF 28-JUN-2002; 2002WO-EP007159.
XX
XX
PR 29-JUN-2001; 2001US-0302133P.
XX
XX
PR (NOVS ) NOVARTIS AG.
XX
PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.
XX
XX
PI Herring CT, Langford G, Quinn G, Scobie L;
XX
XX
DR WPI; 2003-210279/20.
XX
XX
PT An isolated polypeptide useful for screening pigs for the presence of a
PT porcine endogenous retrovirus (PERV) and providing a pig modified not to
PT express a PERV.
XX
XX
PS Claim 1; Page 57-58; 63pp; English.
XX
XX
CC The present sequence is that of a claimed flanking sequence of porcine
CC endogenous retrovirus (PERV). Flanking sequences are DNA sequences
CC surrounding a PERV which represent a unique molecular signature that can
CC be used to characterise PERV integration sites and identify PERVs. The
CC present sequence is one of eight claimed PERV polynucleotides (see
CC ABZ57861-68) useful for screening a pig for the presence of PERV. The
CC screening provides for the elimination of donors with known proviruses. A
CC pig modified not to express a selected PERV is obtained by identifying a
CC PERV using the claimed polynucleotides, and knock-out or inactivation of
CC the PERV by homologous recombination. The organs, tissues and cells of
CC the modified pig are suitable for use in xerotransplantation
XX
SQ Sequence 2860 BP; 691 A; 492 C; 533 G; 1144 T; 0 U; 0 Other;

Query Match 1.9%; Score 51; DB 8; Length 2860;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 TAATACGACTCTACTATAGGCGACGCGTGTGTCGACGCGCGGCTGTCTGCGTGC 57
DB 2 TAATACGACTCTACTATAGGCGACGCGTGTGTCGACGCGCGGCTGTCTGCGTGC 52

RESULT 12
ABZ57865/c
-ID ABZ57865 standard; DNA; 1976 BP.
XX
XX
AC ABZ57865;
XX
XX
DT 14-APR-2003 (first entry)
XX
DE Porcine endogenous retrovirus flanking sequence.
XX
KW PERV; knockout animal; pig; xerotransplantation; organ transplant; ss.
XX
OS Porcine endogenous retrovirus.
XX
FH Key Location/Qualifiers
FT misc_feature 1..54
FT /tag= a
FT /note= "vector"
FT primer_bind 55..75
FT /tag= b
FT LTR 76..504
FT /tag= c
FT primer_bind 486..504
FT /tag= d
FT primer_bind 1676..1695
FT /tag= e
FT misc_feature 1704..1976
FT /tag= f
FT /note= "vector"
FT primer_bind 1704..1733
```


THIS PAGE LEFT BLANK

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2005, 21:32:08 ; Search time 451 Seconds
(without alignments)
9759.611 Million cell updates/sec

Title: US-10-751-550-1

Perfect score: 2690

Sequence: 1 tctagataatacagactcact.....ccataactagtagacaccatgg 2690

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	2.1	7218	1	US-08-232-463-14
2	54	2.0	1141	4	US-09-806-7088-22
3	53.4	2.0	51259	3	US-08-781-891-209
4	53.4	2.0	51259	4	US-09-618-166-209
5	51.8	1.9	3718	3	US-09-424-283-6
6	50.8	1.9	7218	1	US-08-232-463-14
7	49.2	1.8	1926	3	US-09-249-585A-2
8	49.2	1.8	1926	4	US-09-410-399-3
9	49.2	1.8	2580	3	US-09-050-863-2
10	49.2	1.8	2580	3	US-09-359-081-2
11	49.2	1.8	5452	2	US-09-130-114-1
12	49.2	1.8	8705	4	US-09-647-344A-14
13	49.2	1.8	9600	3	US-08-910-647-1
14	49.2	1.8	10596	1	US-07-884-811-15
15	49.2	1.8	10596	1	US-07-885-971-15
16	49.2	1.8	10596	1	US-08-087-782A-15
17	49.2	1.8	10596	1	US-08-194-088B-15
18	49.2	1.8	10596	1	US-08-194-087-15
19	49.2	1.8	10596	5	PCT-US93-04648-15
20	49.2	1.8	16080	4	US-09-724-566A-48
21	49.2	1.8	16080	4	US-09-471-669A-48
22	49.2	1.8	289	3	US-09-007-005-17
23	48.4	1.8	289	3	US-09-244-796-17
24	48.4	1.8	767677	4	US-09-949-016-12147
25	47.8	1.8	767677	4	US-09-949-016-12147
26	47.8	1.8	767677	4	US-09-949-016-17361
27	47.6	1.8	648	4	US-09-598-401C-37

28	47.4	1.8	288	4	US-09-598-401C-38	Sequence 38, Appl
29	47.4	1.8	1924	3	US-09-424-283-5	Sequence 5, Appli
30	47.2	1.8	5185	4	US-09-976-594-640	Sequence 640, App
31	47	1.7	48	3	US-08-913-014A-18	Sequence 18, Appl
32	47	1.7	48	4	US-09-402-532-35	Sequence 35, Appl
33	47	1.7	48	4	US-09-653-285-18	Sequence 18, Appl
34	47	1.7	921	3	US-09-377-648-4	Sequence 4, Appli
35	46.4	1.7	28806	4	US-09-949-016-13217	Sequence 13217, A
36	46	1.7	1141	4	US-09-806-7088-22	Sequence 22, Appl
37	45	1.7	12695	4	US-09-949-016-16775	Sequence 16775, A
38	44.4	1.7	47	2	US-08-582-562A-53	Sequence 53, Appl
39	44.4	1.7	47	2	US-08-778-494B-53	Sequence 53, Appl
40	44.4	1.7	554	4	US-09-489-039A-1758	Sequence 1758, Ap
41	44.4	1.7	570	4	US-09-489-039A-1758	Sequence 1676, Ap
42	44.4	1.7	150394	4	US-09-949-016-13042	Sequence 13042, A
43	44.2	1.6	2614	4	US-09-004-056-1	Sequence 1, Appli
44	44.2	1.6	6078	3	US-09-173-914-1	Sequence 1, Appli
45	44	1.6	16442	3	US-08-781-891-208	Sequence 208, App

ALIGNMENTS

RESULT 1
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZ9pt-Fls
US-08-232-463-14

Query Match 2.1%; Score 56; DB 1; Length 7218;


```
;; APPLICANT: DORNER, F.
;; APPLICANT: SCHEIFLINGER, F.
;; APPLICANT: FALKNER, F. G.
;; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
;; NUMBER OF SEQUENCES: 52
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Foley & Lardner
;; STREET: 1800 Diagonal Road, Suite 500
;; CITY: Alexandria
;; STATE: VA
;; COUNTRY: USA
;; ZIP: 22313-0299
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/232,463
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/07/935,313
;; FILING DATE:
;; APPLICATION NUMBER: EP 91 114 300.6
;; FILING DATE: 26-AUG-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: BENT, Stephen A.
;; REGISTRATION NUMBER: 29,768
;; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703)836-9300
;; TELEFAX: (703)683-4109
;; TELEX: 899149
;; INFORMATION FOR SEQ ID NO: 14:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 7218 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; IMMEDIATE SOURCE:
;; CLONE: pTZ9pt-F15
;; US-08-232-463-14

Query Match 1.9%; Score 50.8; DB 1; Length 7218;
Best Local Similarity 6.0%; Pred. No. 0.0019;
Matches 25; Conservative 217; Mismatches 174; Indels 0; Gaps 0;

QY 1276 TTGAGCATGTTTATATTAGGAGGAGCTTTTACCTGTTGAGTAAATCTGAACGCTAAT 1335
Db 1472 TCTATGCAAGTAGTTAAAGAGATAGAAGAAATTTGGTACRRRRRRRRRRRRRRRRRR 1413
QY 1336 AATCGCTTAAGCAAAATATTTTACCTGTTGATTCTAACAAATTTGATGGACAAT 1395
Db 1412 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1353
QY 1396 ATTGATGAGTGACTGACAAATGATTGAAGGCTTTAAAGAGATTCAGAGGATAATCT 1455
Db 1352 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1293
QY 1456 ACAATAAAATGTAAGAGAAAGCAATCAAGTGTGAGATCTGGTGTGAAGACTATT 1515
Db 1292 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1233
QY 1516 TGCCTCTGGGGTAAAGACACAACTTAGTAGTGGCTCAAAATTTGGAGGCCCA 1575
Db 1232 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1173
QY 1576 TGAAGATTGTTAAAGTAATTTGTTTGGATTGACGGAGGCAATTCGAAGTGATCATCTAC 1635
Db 1172 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1113
QY 1636 CTAGAGCTCTCAATGGGAGGTGCTCGAAGACATATTACCATGTGTATGGCAAGAT 1691

;; APPLICANT: DORNER, F.
;; APPLICANT: SCHEIFLINGER, F.
;; APPLICANT: FALKNER, F. G.
;; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
;; NUMBER OF SEQUENCES: 52
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Foley & Lardner
;; STREET: 1800 Diagonal Road, Suite 500
;; CITY: Alexandria
;; STATE: VA
;; COUNTRY: USA
;; ZIP: 22313-0299
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/232,463
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/07/935,313
;; FILING DATE:
;; APPLICATION NUMBER: EP 91 114 300.6
;; FILING DATE: 26-AUG-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: BENT, Stephen A.
;; REGISTRATION NUMBER: 29,768
;; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703)836-9300
;; TELEFAX: (703)683-4109
;; TELEX: 899149
;; INFORMATION FOR SEQ ID NO: 14:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 7218 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; IMMEDIATE SOURCE:
;; CLONE: pTZ9pt-F15
;; US-08-232-463-14

Query Match 1.9%; Score 50.8; DB 1; Length 7218;
Best Local Similarity 6.0%; Pred. No. 0.0019;
Matches 25; Conservative 217; Mismatches 174; Indels 0; Gaps 0;

QY 1276 TTGAGCATGTTTATATTAGGAGGAGCTTTTACCTGTTGAGTAAATCTGAACGCTAAT 1335
Db 1472 TCTATGCAAGTAGTTAAAGAGATAGAAGAAATTTGGTACRRRRRRRRRRRRRRRRRR 1413
QY 1336 AATCGCTTAAGCAAAATATTTTACCTGTTGATTCTAACAAATTTGATGGACAAT 1395
Db 1412 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1353
QY 1396 ATTGATGAGTGACTGACAAATGATTGAAGGCTTTAAAGAGATTCAGAGGATAATCT 1455
Db 1352 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1293
QY 1456 ACAATAAAATGTAAGAGAAAGCAATCAAGTGTGAGATCTGGTGTGAAGACTATT 1515
Db 1292 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1233
QY 1516 TGCCTCTGGGGTAAAGACACAACTTAGTAGTGGCTCAAAATTTGGAGGCCCA 1575
Db 1232 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1173
QY 1576 TGAAGATTGTTAAAGTAATTTGTTTGGATTGACGGAGGCAATTCGAAGTGATCATCTAC 1635
Db 1172 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1113
QY 1636 CTAGAGCTCTCAATGGGAGGTGCTCGAAGACATATTACCATGTGTATGGCAAGAT 1691

Db 1112 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1057

RESULT 7
US-09-249-585A-2/c
; Sequence 2, Application US/09249585A
; Patent No. 6417002
; GENERAL INFORMATION:
; APPLICANT: Hoxlick, Robert
; TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISOMES
; FILE REFERENCE: 0867/0D905
; CURRENT APPLICATION NUMBER: US/09/249,585A
; CURRENT FILING DATE: 1999-02-11
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1926
; TYPE: DNA
; ORGANISM: Epstein Barr Virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1926)
; OTHER INFORMATION: coding strand of EBNA-1 DNA
; US-09-249-585A-2

Query Match 1.8%; Score 49.2; DB 3; Length 1926;
Best Local Similarity 57.0%; Pred. No. 0.0024;
Matches 90; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 329 CTCCTGTAGTATACACTGCGCTCCAGCTCATGCTCGCAAGCTCCAGGGCGCGCG 388
Db 466 CTCCTGCCCTCCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 407
QY 389 GCAGTGTCTCAACACTTTTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 448
Db 406 CCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 347
QY 449 ACAGCTCTCTCAACACTTTTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 486
Db 346 CTGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 309

RESULT 8
US-09-410-399-3/c
; Sequence 3, Application US/09410399
; Patent No. 6482587
; GENERAL INFORMATION:
; APPLICANT: Robertson, Erle S.
; APPLICANT: Cotter, Murray A.
; TITLE OF INVENTION: Methods to Inhibit or Enhance the Binding of Viral DNA
; TITLE OF INVENTION: to Genomic Host DNA
; FILE REFERENCE: UM-03778
; CURRENT APPLICATION NUMBER: US/09/410,399
; CURRENT FILING DATE: 1999-10-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1926
; TYPE: DNA
; ORGANISM: Epstein-Barr virus
; US-09-410-399-3

Query Match 1.8%; Score 49.2; DB 4; Length 1926;
Best Local Similarity 57.0%; Pred. No. 0.0024;
Matches 90; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 329 CTCCTGTAGTATACACTGCGCTCCAGCTCATGCTCGCAAGCTCCAGGGCGCGCG 388
Db 466 CTCCTGCCCTCCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 407
QY 389 GCAGTGTCTCAACACTTTTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 448
```



```
; SEQ ID NO 1
; LENGTH: 5452
; TYPE: DNA
; ORGANISM: VEBNA
; US-09-130-114-1

Query Match      1.8%; Score 49.2; DB 2; Length 5452;
Best Local Similarity 57.0%; Pred. No. 0.0046;
Matches 90; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 329 CTCCTGTAGTATCACCTGCGTCCAGCTTCAGCTCATGCTCCAGAGCTCCAGGCGGCGCCG 388
Db 1956 CTCCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2015

QY 389 GCAGTGTCCAAACACTTTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 448
Db 2016 CCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2075

QY 449 ACGCAGCTTCTCCACCTTTTACTCTTTTCTTTCTTTCTTT 486
Db 2076 CTGCCCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2113

RESULT 12
US-09-647-344A-14
; Sequence 14, Application US/09647344A
; Patent No. 6586180
; GENERAL INFORMATION:
; APPLICANT: Ruffner, Duane E.
; APPLICANT: Pierce, Michael L.
; APPLICANT: Chen, Zhidong
; TITLE OF INVENTION: Directed Antisense Libraries
; FILE REFERENCE: T6678 PCT/US
; CURRENT APPLICATION NUMBER: US/09/647,344A
; PRIORITY FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: PCT/US99/06742
; PRIOR FILING DATE: 1999-03-28
; NUMBER OF SEQ ID NOS: 50
; SEQ ID NO 14
; LENGTH: 8705
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pShuttle
; US-09-647-344A-14

Query Match      1.8%; Score 49.2; DB 4; Length 8705;
Best Local Similarity 57.0%; Pred. No. 0.0062;
Matches 90; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 329 CTCCTGTAGTATCACCTGCGTCCAGCTTCAGCTCATGCTCCAGAGCTCCAGGCGGCGCCG 388
Db 7821 CTCCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7880

QY 389 GCAGTGTCCAAACACTTTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 448
Db 7881 CCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7940

QY 449 ACGCAGCTTCTCCACCTTTTACTCTTTTCTTTCTTTCTTT 486
Db 7941 CTGCCCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7978

RESULT 13
US-08-910-647-1/c
; Sequence 1, Application US/08910647
; Patent No. 6251433
; GENERAL INFORMATION:
; APPLICANT: Zuckermann et al.
; TITLE OF INVENTION: Compositions and Methods for
; TITLE OF INVENTION: Polynucleotide Delivery
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:

; SEQ ID NO 1
; LENGTH: 5452
; TYPE: DNA
; ORGANISM: VEBNA
; US-09-130-114-1

Query Match      1.8%; Score 49.2; DB 3; Length 9600;
Best Local Similarity 57.0%; Pred. No. 0.0046;
Matches 90; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 329 CTCCTGTAGTATCACCTGCGTCCAGCTTCAGCTCATGCTCCAGAGCTCCAGGCGGCGCCG 388
Db 895 CTCCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 836

QY 389 GCAGTGTCCAAACACTTTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 448
Db 835 CCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 776

QY 449 ACGCAGCTTCTCCACCTTTTACTCTTTTCTTTCTTTCTTT 486
Db 775 CTGCCCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 738

RESULT 14
US-09-620-925-1/c
; Sequence 1, Application US/09620925
; Patent No. 6468986
; GENERAL INFORMATION:
; APPLICANT: Zuckermann et al.
; TITLE OF INVENTION: Compositions and Methods for
; TITLE OF INVENTION: Polynucleotide Delivery
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94508-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/620,925
; FILING DATE: 21-Jul-2000
; CLASSIFICATION: <Unknown>
```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/910,647
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Fujita, Sharon M.
; REGISTRATION NUMBER: 38,459
; REFERENCE/DOCKET NUMBER: 1218.002
; TELEPHONE: (510) 923-2706
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9600 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-620-925-1

Query Match          1.8%; Score 49.2; DB 3; Length 9600;
Best Local Similarity 57.0%; Pred. No. 0.0066;
Matches 90; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 329 CTCCTGTAGTATCACCTGCGTCCAGCTCAGTCTCGCAAGCCTCCAGGGCGGCCG 388
Db      |||||
895 CTCCTGCCCTCTCTCTGCTCTGCTGCGCCCTCTCTGCTCTGCTGCGCCCTCTCTG 836
QY 389 GCAGTGTCTCCACACTTTGCGCTCTCTCTACAGTCTCTTCCACATGCACTGCTCG 448
Db      |||||
835 CCCCTCCTCTCTCTGCTGCGCCCTCTCTGCTCTCTGCTGCGCCCTCTCTGCGCCCT 776
QY 449 ACGCACCTTCTCCACCTTTTACTCTTTTCTTTCTTTCTTTCTTTCTTTCTTT 486
Db      |||||
775 CTGCCCCCTCTCTCTGCTCTGCTGCGCCCTCTCTCTCTCTCTCTCTCTCTCTCT 738

; SEQUENCE CHARACTERISTICS:
; LENGTH: 10596 bases
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-07-884-811-15

Query Match          1.8%; Score 49.2; DB 1; Length 10596;
Best Local Similarity 57.0%; Pred. No. 0.0071;
Matches 90; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 329 CTCCTGTAGTATCACCTGCGTCCAGCTCAGTCTCGCAAGCCTCCAGGGCGGCCG 388
Db      |||||
2429 CTCCTGCCCTCTCTCTGCTCTGCTGCGCCCTCTCTGCTCTGCTGCGCCCTCTCTG 2370
QY 389 GCAGTGTCTCCACACTTTGCGCTCTCTCTACAGTCTCTTCCACATGCACTGCTCG 448
Db      |||||
2369 CCCCTCCTCTCTCTGCTGCGCCCTCTCTGCTCTCTGCTGCGCCCTCTCTGCGCCCT 2310
QY 449 ACGCACCTTCTCCACCTTTTACTCTTTTCTTTCTTTCTTTCTTTCTTTCTTT 486
Db      |||||
2309 CTGCCCCCTCTCTCTGCTCTGCTGCGCCCTCTCTCTCTCTCTCTCTCTCTCTCT 2272

Search completed: May 9, 2005, 02:22:57
Job time : 453 secs
```

```

RESULT 15
US-07-884-811-15/c
; Sequence 15, Application US/07884811
; Patent No. 5316921
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.
; TITLE OF INVENTION: SINGLE-CHAIN HEPATOCYTE GROWTH FACTOR VARIANTS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/884,811
; FILING DATE: 19920518
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: 755.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3216
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 15:
```

THIS PAGE LEFT BLANK

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2005, 21:38:04 ; Search time 8593 Seconds
(without alignments)
11915.848 Million cell updates/sec

Title: US-10-751-550-1
Perfect score: 2690
Sequence: 1 tctagataatagactact.....ccataactagtacaccattgg 2690

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	235	8.7	670	9	CL922874 OA_Aba002
C 2	226.4	8.4	629	9	CL914408 OA_Aba001
C 3	218.8	8.1	1020	9	CL947599 OeIFSB005
C 4	212.4	7.9	1293	9	CL982445 OeIFSC047
C 5	210.6	7.8	732	9	CL936467 OA_Aba004
C 6	209	7.8	729	9	CL935937 OA_Aba004
C 7	207.8	7.7	718	9	CL907647 OA_Aba000
C 8	203.2	7.6	757	9	CL927201 OA_Aba003
C 9	202	7.5	718	9	CL910742 OA_Aba001
C 10	194	7.2	4152	9	CL958665 OeIFCC001
C 11	183.4	6.8	669	9	CL916614 OA_Aba001
C 12	176	6.5	707	9	CL922867 OA_Aba002
C 13	175.8	6.5	678	9	CL908388 OA_Aba000
C 14	175	6.5	902	8	BZ797683 PUGBP407D
C 15	175	6.5	949	8	BZ797681 PUGBP407B
C 16	172	6.4	757	9	CL918926 OA_Aba002
C 17	171.6	6.4	722	9	CL9149354 104_330.1
C 18	166.8	6.2	813	9	CL907539 OA_Aba000
C 19	164.4	6.1	850	9	CL941236 OA_Aba005
C 20	164.2	6.1	779	8	BZ747430 PUDCK347D
C 21	164.2	6.1	836	8	CC404666 FUHUK337D
C 22	163.2	6.1	660	9	CL923944 OA_Aba002
C 23	162	6.0	909	8	CC436809 FUHQ0267D
C 24	161.4	6.0	801	9	CL903983 OA_Aba000

25	161	6.0	749	9	CG867026	CG867026 ZMMBB035
26	160	5.9	729	9	CG828183	CG828183 ZMMBB017
27	160	5.9	773	9	CL917811	CL917811 OA_Aba002
28	159.8	5.9	650	8	CC374857	CC374857 PUEJ53TD
29	159.4	5.9	722	8	CC008427	CC008427 PUDCK10TD
30	158.8	5.9	884	9	CG247085	CG247085 OG0AT79TH
31	158.4	5.9	697	9	CL933353	CL933353 OA_Aba004
32	158.2	5.9	966	8	BZ819943	BZ819943 PUGDI31TD
33	157.4	5.9	686	9	CL938390	CL938390 OA_Aba004
34	157.2	5.9	770	9	CL941728	CL941728 OA_Aba005
35	156.8	5.8	711	9	CL922324	CL922324 OA_Aba002
36	156.8	5.8	1028	9	CG969919	CG969919 ZUAB338TV
37	156.2	5.8	726	8	AQ274269	AQ274269 nbxb0034K
38	156	5.8	681	9	CL933514	CL933514 OA_Aba004
39	155.6	5.8	790	9	CC734662	CC734662 OGVA46TH
40	155.6	5.8	838	9	CG002979	CG002979 ZUAD61LTH
41	155.6	5.8	842	8	BZ795811	BZ795811 PUFDS84TE
42	155.6	5.8	927	8	CC347166	CC347166 OG0BH58TV
43	155.4	5.8	627	9	CG937241	CG937241 ZMMBB022
44	155.2	5.8	775	9	CL906326	CL906326 OA_Aba000
45	154.6	5.7	655	8	AQ577400	AQ577400 nbxb0091A

ALIGNMENTS

RESULT 1
CL922874/c
LOCUS CL922874 670 bp DNA linear GSS 14-SEP-2004
DEFINITION OA_Aba0028B23.r OA_Aba Oryza australiensis genomic clone
ACCESSION OA_Aba0028B23 3', Genomic survey sequence.
VERSION CL922874.1 GI:52039943
KEYWORDS GSS.
SOURCE Oryza australiensis
ORGANISM Oryza australiensis

REFERENCE
AUTHORS Kim,H., Yu,Y., Stum,D., Yost,D., Rao,K., Luo,M., Jetty,R.,
Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.
TITLE OMAP Project
JOURNAL Unpublished (2004)
COMMENT Contact: Rod A. Wing
Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu
PCR PRIMERS
FORWARD: TAC TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Plate: 0028 row: B column: 23
Seq primer: CAC TCA TTA GGC ACC CCA
Class: BAC ends.

FEATURES
source
1..670
/organism="Oryza australiensis"
/mol_type="genomic DNA"
/db_xref="taxon:4532"
/clone="OA_Aba0028B23"
/tissue_type="young leaves"
/lab_host="DH10B Ti phage resistant"
/clone_lib="OA_Aba"
/note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"

ORIGIN
Query Match 8.7%; Score 235; DB 9; Length 670;
Best Local Similarity 68.0%; Pred. No. 2.2e-52;
Matches 387; Conservative 0; Mismatches 175; Indels 7; Gaps 4;

Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.

FEATURES
source
1..1020
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Expressed Library"
/note="Oryza sativa exon trapped genomic sequences"

ORIGIN

Query Match 8.1%; Score 218.8; DB 9; Length 1020;
Best Local Similarity 70.9%; Pred. No. 6.6e-48;
Matches 343; Conservative 0; Mismatches 117; Indels 24; Gaps 3;
QY 1232 GAGCTCATCTTATATAGCATGCTGATACCAAAATTAATCTCTTTGAGCATGTTTATA 1291
DB 533 GGGCTCATCTATATAGCATGCTGCTACTAAAGTTACTCTTTGAGCTTGTGTATG 592
QY 1292 TTTAGGAGGACGTTTACCTGTTGAGGTAAATCTGAACGCTAATAATCGGCTAAGCAAA 1351
DB 593 GTCAGAGAGCCGTTTCCCGTTGAGGTAAATCTGGACGCTTATAGATTGGCTTAACAAA 652
QY 1352 ATAAATTTATCACCTGTGTATCTCAAAATTTGATGATGGACAATATTGATGAGGTGACTG 1411
DB 653 ATGACCTTGATGCTGTTGATTACCAAAATTTGATGTTGGACAATATTGATGAGGTAACCG 712
QY 1412 ACAATGATTTGAGGCTTTAAAGGAGATTGAGAAGAT-----AAATCT 1455
DB 713 ACAAAAGATTGAAGAGCTTTGAAGGAGATTGAGAAGATTAAGCTTCGGGTTGCTAAAGCTT 772
QY 1456 ACAATAAAATGTAAGAAGAAAGCAATTCAAAGTGTGAGATCTGGTGTGGAAGACTATTT 1515
DB 773 ACAACAGAAAGTACGAGAAATATTTTCAGATTGGTGAGCTTTGTTGGAAGACGATCT 832
QY 1516 TGCCTCTTGGGGTAAAGACAAACAGTTTATAGTGGCCCTCAAAATTTGGGAGGCCCA 1575
DB 833 TGCCTATTGGGACGAAAG-CAACAAGTTTGGTAAATGTCACCAAGTTGGGAAGGACCC 891
QY 1576 TCGAAGATTGTTAAAGTAATGTTTGGAT-----TGACGGAGCATTTCAAGTGAT 1628
DB 892 TATAGAGTTGTTAAAGTCAATTTTGGGAATTTCTTACATGCTAGACGCTGAAGGGAGAT 951
QY 1629 CATCTACCTAGAGCTCTCAATGGGAGTGTCTCGAAGACATATTACCCTGTGTATGGCAA 1688
DB 952 CATTTGCCCTAAAGCTATCAACGGAAGATCTTGAAGAAATTTATCCAAAGTTTGGCAA 1011
QY 1689 GATG 1692
DB 1012 GACG 1015

RESULT 4
CL982445
LOCUS
DEFINITION
OeIPSC047866 Oryza sativa Expressed Library Oryza sativa (indica
cultivar-group) genomic, genomic survey sequence.

ACCESSION
CL982445
VERSION
CL982445.1 GI:52419372
KEYWORDS
GSS.
SOURCE
Oryza sativa (indica cultivar-group)
ORGANISM
Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
1 (bases 1 to 1293)
AUTHORS
Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,

Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
Wong, G.K.S., Deng, X.W. and Wang, J.
An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
Unpublished (2004)
Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.

FEATURES
source
1..1293
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Expressed Library"
/note="Oryza sativa exon trapped genomic sequences"

ORIGIN

Query Match 7.9%; Score 212.4; DB 9; Length 1293;
Best Local Similarity 70.0%; Pred. No. 4e-46;
Matches 339; Conservative 0; Mismatches 121; Indels 24; Gaps 3;
QY 1232 GAGCTCATCTTATATAGCATGCTGATACCAAAATTAATCTCTTTGAGCATGTTTATA 1291
DB 806 GGGCTCATCTATATAGCATGCTGCTACTAAAGTTACTCTTTGAGCTTGTGTATG 865
QY 1292 TTTAGGAGGACGTTTACCTGTTGAGGTAAATCTGAACGCTAATAATCGGCTAAGCAAA 1351
DB 866 GTCAGAGAGCCGTTTCCCGTTGAGGTAAATCTGGACGCTTATAGATTGGCTTAACAAA 925
QY 1352 ATAAATTTATCACCTGTGTATCTCAAAATTTGATGATGGACAATATTGATGAGGTGACTG 1411
DB 926 ATGACCTTGATGCTGTTGATTACCAAAATTTGATGTTGGACAATATTGATGAGGTAACCG 985
QY 1412 ACAATGATTTGAGGCTTTAAAGGAGATTGAGAAGATTAAGCTTCGGGTTGCTAAAGCTT 1455
DB 986 ACAACAGATTGAAGAGCTTTGAAGGAGATTGAGAAGATTAAGCTTCGGGTTGCTAAAGCTT 1045
QY 1456 ACAATAAAATGTAAGAAGAAAGCAATTCAAAGTGTGAGATCTGGTGTGGAAGACTATTT 1515
DB 1046 ACAACAGAAAGTACGAGAAATATTTTCAGATTGGTGAGTTTGTGGAAGACGATCT 1105
QY 1516 TGCCTCTTGGGGTAAAGACAAACAGTTTATAGTGGCCCTCAAAATTTGGGAGGCCCA 1575
DB 1106 TGCCTATTGGGACGAAAG-CAACAAGTTTGGTAAATGTCACCAAGTTGGGAAGGACCC 1164
QY 1576 TCGAAGATTGTTAAAGTAATGTTTGGAT-----TGACGGAGCATTTCAAGTGAT 1628
DB 1165 TATAGAGTTAGTCCAAAGTCAATTTTCGGGAATTTTACATGCTAGACGCTGAAGGGAGAT 1224
QY 1629 CATCTACCTAGAGCTCTCAATGGGAGTGTCTCGAAGACATATTACCCTGTGTATGGCAA 1688
DB 1225 CATTTGCCCTAAAGTATCAATGGGAGATTTGAAGAAATTTATCCAAAGTTTGGCAA 1284
QY 1689 GATG 1692
DB 1285 GATG 1288

RESULT 5
CL936467
LOCUS
DEFINITION
OA_ABa0047D02.r OA_ABa Oryza australiensis genomic clone
ACCESSION
CL936467
VERSION
CL936467.1 GI:52067031
KEYWORDS
GSS.
SOURCE
Oryza australiensis

CL936467
OA_ABa0047D02.r OA_ABa Oryza australiensis genomic clone
CL936467
CL936467.1 GI:52067031
GSS.
Oryza australiensis

```
ORGANISM Oryza australiensis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 732)
AUTHORS Kim,H., Yu,Y., Stum,D., Yost,D., Rao,K., Luo,M., Jetty,R.,
Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.
TITLE OMAP Project
JOURNAL Unpublished (2004)
COMMENT Contact: Rod A. Wing
Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu
PCR Primers
FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Plate: 0047 row: D column: 02
Seq primer: CAC TCA TTA GGC ACC CCA
Class: BAC ends.
FEATURES             Location/Qualifiers
     source
       1..732
         /organism="Oryza australiensis"
         /mol_type="genomic DNA"
         /db_xref="taxon:4532"
         /clone="OA_Aba0047D02"
         /tissue_type="young leaves"
         /lab_host="DH10B T1 phage resistant"
         /clone_lib="OA_Aba"
         /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"
ORIGIN
Query Match      7.8%; Score 210.6; DB 9; Length 732;
Best Local Similarity 73.2%; Pred. No. 1.1e-45;
Matches 311; Conservative 0; Mismatches 109; Indels 5; Gaps 3;
QY 659 CCCACACTTAGCCCTTGCTCATCTCGAGTAAGTTCAAGACATGAGTGAACATCT 718
Db 1 CCCACACTTAGCCCTTGCTCATCTCGAGTAAGTTCAAGTACTAAGTGGATCATGA 60
QY 719 CCTCAATGGTACGATGCCCTGCAATAAGTTATTCGAAGCCTCACCTATACATGCAACT 778
Db 61 CTTCTTTTGGTACAAGTTGGACATACAAGTTATTCGAACATACCTATATCTGTGAAT 120
QY 779 TTGAAGTGCTACACGCCATCTGGGTGGTTGAGAAATGGAAACAGATCAGAAATCCAGTC 838
Db 121 TGGATGTGTACCATGTTATCTCGAGTTATTGAAGAATGGAACAGCTAGGTTAAAGTT 180
QY 839 ATCTTTACCTCTCTGCTTAGATAACTTGGGTTTTTTGTAAGGTTTTCAAAATTT---AAAA 895
Db 181 ACTTTAATCTCTGCTTAATGACTTTGAGGTTTTCAATAGTTTTCAAAATGAAAAA 240
QY 896 CATAGTCTGCTCCTCAATGATTTCTCATATPAGCTCAATGTTGATGGTTCTCACCAA 955
Db 241 CATAGCTTTACTCCTAGGATGATTTCTTAGGTCACTCAATGTTGTTGATTTCTCACCAA 300
QY 956 GGCAATGTTTGGCTCTTTCATCCCTACTTCTAATATTTCTTTTGGAGCTTAGGGTAG 1015
Db 301 GGCATGATTTGGCTCTTTTC-FCCCTACTTCTTATAAGGCTATATGGGGCTCAAGATAG 359
QY 1016 GGAATGAAAGGAAGCATACTTTGCATTTGCATATGTTTACTTAAGTCAAAACCAAACTCGAG 1075
Db 360 GGAA-AATCATGAGGCATACTTGCATCAATATTTGTAAAGTCAAAAGTATATCCAG 418
QY 1076 GAGAA 1080
Db 419 AAGAA 423
RESULT 6
CL935937/c
```

```
LOCUS CL935937 729 bp DNA linear GSS 14-SEP-2004
DEFINITION OA_Aba0046H01.r OA_Aba Oryza australiensis genomic clone
ACCESSION CL935937
VERSION CL935937.1 GI:52065973
KEYWORDS Oryza australiensis
SOURCE Oryza australiensis
ORGANISM Oryza australiensis
REFERENCE 1 (bases 1 to 729)
AUTHORS Kim,H., Yu,Y., Stum,D., Yost,D., Rao,K., Luo,M., Jetty,R.,
Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.
TITLE OMAP Project
JOURNAL Unpublished (2004)
COMMENT Contact: Rod A. Wing
Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu
PCR Primers
FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Plate: 0046 row: H column: 01
Seq primer: CAC TCA TTA GGC ACC CCA
Class: BAC ends.
FEATURES             Location/Qualifiers
     source
       1..729
         /organism="Oryza australiensis"
         /mol_type="genomic DNA"
         /db_xref="taxon:4532"
         /clone="OA_Aba0046H01"
         /tissue_type="young leaves"
         /lab_host="DH10B T1 phage resistant"
         /clone_lib="OA_Aba"
         /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"
ORIGIN
Query Match      7.8%; Score 209; DB 9; Length 729;
Best Local Similarity 65.1%; Pred. No. 2.9e-45;
Matches 372; Conservative 0; Mismatches 190; Indels 9; Gaps 4;
QY 517 TGTCCTCCCTTACAAATGATAAATCACAAACATCATGGAGCTTGCTAGTTATAAACTTAA 576
Db 590 TATTCACTGCATACAAATAATCACCAACACTCATGGAAGAGTTAGTAAATAACCTTAC 531
QY 577 TTCTAAGTTTGGTGTATTATTGAGTGGATTTTGTCTGAAAGTTGGTGTAGAAATAGG- 635
Db 530 CGCTATGTTGATGTGCATCGTTTGTTCAGTTTATGACGGATTTGACGGCGGATTTTGGC 471
QY 636 AGTTAAGGACGCCCAACAGATCCCCACACTTACCTTGGCTCATCTCGAGTAAAGT 695
Db 470 AGTTAAGGACCGTCAACAAGCTCCCTACACATGCCATTTGCTTGTCTTGAGTAAAGG 411
QY 696 TCAAGGACTAAGTGGACATCTCTCTCAATGGTACGATGCTGCATATAGTTATTCCA 755
Db 410 TTGGAATACTAAGTGGGATCATGATTACTTTTGGTCAATGCTAATATACAAATTTCTCT 351
QY 756 AGCCTCACCTATACATGTGAATCTTTGAAGTGTCTACCAAGCTCTTGGGTGGTTGAGAA 815
Db 350 AGACTTACTGTATCTGTGATCTGAATTTGTCTACCATTTGTATCTTAAGTTATTGAAGA 291
QY 816 ATGGAACAGATCAGAAATCCAGTCATCTTTTACTCTCTTGTCTTAGATAAATCTGGGTTTT 874
Db 290 AGGGAACCAATCTAGGTTTAAAGTACTTTTATTTCTCTGCTTAAATGACTTGGGGTTTT 231
QY 875 -----GTAAGTTTTCAATTTTAAACATAGCTTGTCTCCCAATGATTTCTTCATATA 929
Db 230 AAGAAGTTCTTTTGAAGAAAGAAAAATAATTTTACTCTCTCTGTGATGATTTCTCTCATATC 171
```


Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.

FEATURES

source
1. .4152
Location/Qualifiers
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Express Library"
/note="Oryza sativa exon trapped genomic sequences"

ORIGIN

Query Match 7.2%; Score 194; DB 9; Length 4152;
Best Local Similarity 67.4%; Pred. No. 5.8e-41;
Matches 329; Conservative 0; Mismatches 135; Indels 24; Gaps 3;
QY 1232 GAGCTCATCTTATATAGACATGGTGATACCAAAATTTACTCTTTTGAGCATGTTTATA 1291
DB 3665 GGGCTCATCGTATCTAGACATGGTGCAAAAGGTAAACCAATTTGAATTAGTATATG 3724
QY 1292 TTTAGAGACGTTTACTCTGTTGAGTAAATCTGAACGCTAATAAATCGGCTAAGCAA 1351
DB 3725 GGCAGAGGCCATTTTGCATGTTGAGGTGAATCTGGACGCCCTTAGATTGGCCAGACAA 3784
QY 1352 ATAAATTTATCACTGTTGATCTTCAAAATTTGATGATGACATATTTGATGAGTGACTG 1411
DB 3785 ATAATTTGCGCAGTAGATTTACAACTTAATGATGGACGGAATAGATGAGGTAGTG 3844
QY 1412 ACAATGATTTGAAGGCTTTAAAGGAGATTGAGAGGATAAAT-----CT 1455
DB 3845 ATGAAGATTGAAGCTTTTGAGGAGATTGAGAAAGATTAATTTGAGAGTAGCTAAGCTT 3904
QY 1456 ACAATAAAAATGTAAAGAAAGCAATTCAAAGTGTGAGATCTGSGTGTGGAAGACTATTT 1515
DB 3905 ACAATAAAAGGTGAAGGAAAAATCGTTTCAAATTTGGAGATTTAGTGTGGAACAAATTC 3964
QY 1516 TGCCCTCTGGGGTAAAGACAAACAGTTAGTAACTGAGCTCAAAATTTGGGAGGCCCA 1575
DB 3965 TGCCCTGTT--GGTTCAAAGATATAAATTTGGGAAATGTTCTTCAAATTTGGGAGGCCCA 4023
QY 1576 TGCAAGATTGTTAAAGTAAT-----TGTTTGGATTGACGAGGCAATTTCAAGGTGAT 1628
DB 4024 TATAGATTGTAGAATAGTCCCGGGAATCTTATTTTGTGCAAGTCTACGAGATAT 4083
QY 1629 CATCTACCTAGAGCTCTCAATGGGAGGTGCTCGAAGACATATTTACCCATGTGTATGGCAA 1688
DB 4084 AAGTTGCCCTAGAGCTCTTAATGGGAGATATTTAAAAAGGTACTATCCAAAGTGTGGCAA 4143
QY 1689 GATGTTTA 1696
DB 4144 GATGCTTA 4151

RESULT 11
CL916614/c
LOCUS
DEFINITION
OA_ABa0019F04.r OA_ABa Oryza australiensis genomic clone
OA_ABa0019F04 3', genomic survey sequence.
ACCESSION
CL916614
VERSION
GSS.
KEYWORDS
SOURCE
ORGANISM
Oryza australiensis
Oryza australiensis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
1 (bases 1 to 669)
Kim H., Yu Y., Stum D., Yost D., Rao K., Luo M., Jetty R.,
Kudrna D., Muller C., Hatfield J., Soderlund C. and Wing R.
OWAP Project

JOURNAL
COMMENT

Unpublished (2004)
Contact: Rod A. Wing
Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu
PCR Primers
FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Plate: 0019 row: F column: 04
Seq primer: CAC TCA TTA GGC ACC CCA
Class: BAC ends.

FEATURES

source
1. .669
Location/Qualifiers
/organism="Oryza australiensis"
/mol_type="genomic DNA"
/db_xref="taxon:4532"
/clone="OA_ABa0019F04"
/tissue_type="young leaves"
/lab_host="DH10B Tl phage resistant"
/clone_lib="OA_ABa"
/note="Vector: pGIBAC1; Site_1: HindIII; Site_2: HindIII"

ORIGIN

Query Match 6.8%; Score 183.4; DB 9; Length 669;
Best Local Similarity 65.3%; Pred. No. 2.8e-38;
Matches 331; Conservative 0; Mismatches 171; Indels 5; Gaps 4;
QY 521 CCCCTACAAATGATAAATCACCAAACTCATCGAGCTTCTAGTTATATAAATCTAATTTCT 580
DB 507 CACCTGCATACAAATATTAACCAACACTTGTGGAACAGATTAGTAATAACCTTATCGCT 448
QY 581 AAGTTTGGTGTATTTAGTGGATTTTGTGAAAAGTTGGTGGT-TAGAAATAGGAGTT 639
DB 447 ACATTGATGTTGATCTTTTCTCCGTTTATGCGAGGATTTGATGTCGAATTTTATGATTT 388
QY 640 AAGGACGGCCAAACAGATCCCCACACTTTAGCCCTTTGCTCATCTCCGAGTAAAGTTCAA 699
DB 387 AAGGATCATCAACATGCTCCCCACACTTAGTCTTGTCTGCTCGGGTAAAGCTTAA 328
QY 700 GGACTAAGGTGGAACATCTCTCAAATGTCAGATGCTGTCATATAGTTATTTCCAAGCC 759
DB 327 GTACTAAGGTGATCATGACTTATTTGGTATAATGCTGACATACATGATTTCCAAACC 268
QY 760 TCACCTATACATGTGAAC-TTTGAAGTGTCTACCCAGCCCATCTTGGGTGGTTGAGAAATG 818
DB 267 ATACCTGTACCGGTGAACCTTTTGAAGTGTCTACCATGTCATCTTGGGTAAATGAAGAATG 208
QY 819 GAACAGATCAGATCCAGTCATCTTTACCTCTCTTGTCTTAGATAAATCTTGGTTTTGTAA 878
DB 207 GTACAGCCTTGATCAAATCACCCCTACTTCTGATAAAT-ATCTTAGCGTTTTCTTAT 149
QY 879 GGTTTTCAAATTTAAACATAGTCTTGTCTCCCAATGATTTCTCTCATATAGTCAATGT 938
DB 148 GCTCATAGAAACATAACTTTACTTTTACTCTCCTCAATTTGATTTCTCATATTTGCTCATGT 89
QY 939 GTATGTTTCTTCCAAAGCAAGTGTTCCTCTCTTTTCATCTCTTTCATCTCTTCTTCTTCT 996
DB 88 GTTTTATTCTCACAAGGCTTTGATGTTTGCCTATTCTTATCTCTACCTCTTCTTCTTCTTCT 29
QY 997 TTTTGGGAGCTTAGGTTAGGATGAA 1023
DB 28 TATGTGATGCTTAGGTTAGGATGAAATAA 2

RESULT 12
CL922867/c
LOCUS
DEFINITION
OA_ABa0028B19.r OA_ABa Oryza australiensis genomic clone
OA_ABa0028B19 3', genomic survey sequence.
ACCESSION
CL922867

CL922867 707 bp DNA linear GSS 14-SEP-2004
OA_ABa0028B19.r OA_ABa Oryza australiensis genomic clone
OA_ABa0028B19 3', genomic survey sequence.
CL922867

```

VERSION          CL922867.1  GI:52039929
KEYWORDS
SOURCE           Oryza australiensis
ORGANISM         Oryza australiensis
                  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
                  Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
AUTHORS          Kim,H., Yu,Y., Stum,D., Yost,D., Rao,K., Luo,M., Jetty,R.,
                  Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.
TITLE            OMAP Project
JOURNAL          Unpublished (2004)
COMMENT          Contact: Rod A. Wing
                  Arizona Genomics Institute
                  University of Arizona
                  Forbes Building Room 303, Tucson, AZ 85721-0036, USA
                  Tel: 520 626 9595
                  Fax: 520 621 1259
                  Email: http://genome.arizona.edu
                  PCR Primers
                  FORWARD: TAA TAC GAC TCA CTA TAG GG
                  BACKWARD: CAC TCA TTA GGC ACC CCA
                  Plate: 0028 row: B column: 19
                  Seq primer: CAC TCA TTA GGC ACC CCA
                  Class: BAC ends.

FEATURES
source           Location/Qualifiers
                  1..707
                  /organism="Oryza australiensis"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:4532"
                  /clone="OA_ABa0028819"
                  /tissue_type="young leaves"
                  /lab_host="DH10B T1 phage resistant"
                  /clone_lib="OA_ABA"
                  /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"

ORIGIN
Query Match      6.5%; Score 176; DB 9; Length 707;
Best Local Similarity 65.3%; Pred. No. 3e-36; Indels 9; Gaps 5;
Matches 337; Conservative 0; Mismatches 170;

QY 521 CCCTTCAAAATGATAATCACCACCACTCATGGAGCTTCTAGTTATATAAATCTTAATCT 580
DB 518 CACCTGCATACAAATATCACCAGCACTTGTGGAAGAAATAGTAATAAATCACTACTGCT 459
QY 581 AGTTTGGTGTATTTAGTGGATTTGTGGAAGTTGGTGGTAGAAATAGG--AGT 638
DB 458 ATGTGATGTTATCTTTTATCTATTTATGATGAGGATTCAGAGTCGGATTTTAGTACTT 399
QY 639 TAAGGACCGCCAAACAGATCCCCACACTTAGCCCTTTCCTCATCTCCAGTAAAGTTCA 698
DB 398 CAAGGACTTTTAACATCTCCCTACACCTTGGCCATTCCTGCTCCAGTAAAGTTA 339
QY 699 AGACTAAGTGGAGCAATCTCTCAAATGGTAGATGCTCGCATATPAAGTTATTCGAAGC 758
DB 338 AGTACTAAGTGGATCATGACGCTGTTTGGTATATGCTGACATCAACTATATACAAA 279
QY 759 CTCACCTATACATGTGAACCTTTGAGTGTCTACCGCCATCTGGGTGGTTGAGAAATG 818
DB 278 CATATCTGACCTGTGAAACTTGATGTGCTATCATGTTATCTTGTGCTATTTGAAAGATG 219
QY 819 GAACAGATCAGATCCAGTCATCTTTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 878
DB 218 GAACAGCTAGGTTAAGTCACTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 159
QY 879 GGTTCATCAATTTA---AAACATAGTCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 935
DB 158 AATTTTAAACCGGATAAACATAAATCTTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 100
QY 936 TGTGTATGTTCTTCCACCAAGCAATG-TTTTGGCTCTTTTCTATCTCTCTCTCTCTCTCT 992
DB 99 TGTGTTTGAATCTCTACCAAGGCTTGATGTTGTTCTATTTCTTATCTCTCTCTCTCTCT 40
```

```

QY 993 TTCCTTTTGGAGCTTAGGTTAGGTAAGGAATCAAAAGGA 1028
DB 39 GCCTTATGTGGAGCTCAGGCTAGGGAATAAAGAGTA 4

RESULT 13
CL908388/c
LOCUS
DEFINITION      Oryza australiensis genomic clone
                  OA_ABa0007L13.3', genomic survey sequence.
ACCESSION       CL908388
VERSION         CL908388.1  GI:52017267
KEYWORDS        GSS.
SOURCE          Oryza australiensis
ORGANISM        Oryza australiensis
                  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
                  Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
AUTHORS          Kim,H., Yu,Y., Stum,D., Yost,D., Rao,K., Luo,M., Jetty,R.,
                  Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.
TITLE            OMAP Project
JOURNAL          Unpublished (2004)
COMMENT          Contact: Rod A. Wing
                  Arizona Genomics Institute
                  University of Arizona
                  Forbes Building Room 303, Tucson, AZ 85721-0036, USA
                  Tel: 520 626 9595
                  Fax: 520 621 1259
                  Email: http://genome.arizona.edu
                  PCR Primers
                  FORWARD: TAA TAC GAC TCA CTA TAG GG
                  BACKWARD: CAC TCA TTA GGC ACC CCA
                  Plate: 0007 row: L column: 13
                  Seq primer: CAC TCA TTA GGC ACC CCA
                  Class: BAC ends.

FEATURES
source           Location/Qualifiers
                  1..678
                  /organism="Oryza australiensis"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:4532"
                  /clone="OA_ABa0007L13"
                  /tissue_type="young leaves"
                  /lab_host="DH10B T1 phage resistant"
                  /clone_lib="OA_ABA"
                  /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"

ORIGIN
Query Match      6.5%; Score 175.8; DB 9; Length 678;
Best Local Similarity 64.8%; Pred. No. 3.3e-36;
Matches 370; Conservative 0; Mismatches 192; Indels 9; Gaps 7;

QY 515 AATGTCCTTCAAAATGATAAATCACCACCACTCATGGAGCTTCTAGTTATAAATCTCT 574
DB 588 ATTATTCACCTGGATACAAATATTTACCAACACTCGTTGAAAGAGTTAATAATAAACCT 529
QY 575 AATTCATAGTTTGGTGTATTTATTTAGTGGATTTTGTGGAAGTTGGTGGTTAGA-AATA 633
DB 528 ACCACTACGTTGATGAGCATAGTTTGTTCATTTATGAGGATTCGATGTTTTCGATATTA 469
QY 634 GGAGTTAAGGACCGCCAAACAGATCCCCACACTTAGCCCTTTCCTCATCTCCGAGTAA 693
DB 468 GTCGTTAAGGACCAATAACAGCTCCCAACACTTTGCTTTTGTTCATCTCTCTGTTAA 409
QY 694 GTTCAAG-GACTAAGTGGACATCTCTCTCAATGGTAGCATGCTCGCATATAAGTTATT 752
DB 408 GGCCTGTGATCTAAGTGGATCATGATAAATTTTGGTGTAGAGCAACATACAGTTATT 349
QY 753 CCAAGCCTCAGCTATACATGTGAACCTTTGAAGTGTCTACCAAGCCATCTTGGTGGTTGA 812
DB 348 CTTAGACCTACTCATPACTGATGATCTATATGTGTACCATGCAATTTCTAAGT-ATTGA 290
QY 813 GAAATGGAACAGATCAGAAATCCAGTCATCTTTACCTCTCTTCTTGTAGATAACTTGGTTT 872
```

```
Db      289  AAAATGGAACAG -TCTAGATATAAGCTACTATATTCTGCTGCATTAAATAAAGCTTGGGTT 231
Qy      873  TTGTAAGGTTTC--AAATTTAAACATAGTCTTCTCTCTCAATGATCTCTCATATA 929
Db      230  TTGATAAGTTTTCAAAATAATGTAACATGACTTTACTCTTGTGATGATTTCTCATGTC 171
Qy      930  GCTCAATGTGATGTTCTTCCACCAAGGCAATGTTTGGCTCTCTTTTCATCTCTACTCTAA 989
Db      170  ACTCGAGGTTTGTATCTCTCACCATGGCAAT-ACTTGCACAAATTTTCTTACTTCTAA 112
Qy      990  TATTTCTTTTGTGAGCTTAGGTAAGGTAAGAAAGGAAGCATCTTGCATGTCATATG 1049
Db      111  TAAAGCTTATGTGGAGCTCAAGTAGGAA-AATCATGGCGCATACTTGCATCAAAATATT 53
Qy      1050 TTTACTAAGTCAAACCAATCTGAGGAA 1080
Db      52  TTTTAAAGTCAGATAAAGATTTCAAGAAGAA 22

RESULT 14
LOCUS   BZ797683/c
DEFINITION   PUGBP40TD ZM_0.6_1.0_KB Zea mays genomic clone ZMBTA331G08,
genomic survey sequence.
ACCESSION   BZ797683
VERSION     BZ797683.1 GI:29000291
KEYWORDS    GSS.
SOURCE      Zea mays
ORGANISM    Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 902)
AUTHORS    Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
TITLE      Maize Genomics Consortium
JOURNAL    Unpublished (2003)
COMMENT    Other_GSSs: PUGBP40TB
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.
FEATURES   Location/Qualifiers
            source
            1..902
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /strain="B73"
            /db_xref="taxon:4577"
            /clone="ZMBTA331G08"
            /clone_lib="ZM_0.6_1.0_KB"
            /note="Vector: PCR4-TOPO; Site:1: EcoRI; 0.6-1.0 kb high
            CoT selected genomic DNA library"

ORIGIN
Query Match 6.5%; Score 175; DB 8; Length 902;
Best Local Similarity 64.7%; Pred. No. 5.9e-36;
Matches 319; Conservative 0; Mismatches 150; Indels 24; Gaps 3;

Qy      1232 GAGCTCATCATTTATATAAGCATGGTGATACCAAAATTAATCTCTTTTGAGCATGTTTATA 1291
Db      844  GGGCTCATAGATAATCAAAACTTCATGCCACTAAAGTTTCTCTTTTGAGCTTATGTATG 785
Qy      1292 TTTAGAGGACGCTTTTACCTGTTGAGTTAGTAAATCTGAACCTAATAAATCGGCTAGCAA 1351
Db      784  GAAATGAAGCCATCTTGCCTATAGAAATAAATCTAACCACTATTAGTTTGGCCAGCAA 725
Qy      1352 ATAATTTATCACCTGTTGATTTCTAACAAATTTTGATGATGACAATATTGATGAGGTGACTG 1411
```

```
Db      724  ATAAATTTAGATGCTCGATGAGTACCATGATTTAATGATGACAATATTGACGAGGTGACCG 665
Qy      1412 ACAATGATTTAGAGCTTTAAAGGAGATTTGAGAAGATATAA-----TCT 1455
Db      664  ATAAAGGTTGATGCTTTTGAAGGAAATTTGAAGAGCAAACTTATAGTCGCCAAGGTCT 605
Qy      1456 ACAATAAAAAATGTAAGAGAAAGCATTCAAAGTGTGAGATCTGGTGTGGAAGACTATTT 1515
Db      604  ACAATAGAGTTAAAGACTAAATCATTTTCAGTTGGAGATTTGGTGTGGAAGACAATGT 545
Qy      1516 TGCCTCTTTGGGGTAAAAAGACAAGTTTGTAGTAGTGGCTCAAAATTTGGAGGCCCCA 1575
Db      544  TGCC-ATTGAAGATTAAAGGATTACAAGTTTGAAGTTGACAAGTGGTCTCCAAGTTTGAAGGACT 486
Qy      1576 TCGAAGATTGTTAAAGTAATTTGTTTGGAT-----TCACGGAGGCATTTCAAGTGAT 1628
Db      485  TATAAGGTTGTACATGTAAATCCCTGTAATGCTTATATATGTTGGAAACATTTACAAGGAAAT 426
Qy      1629 CATCTACCTAGAGCTCTCAATGGGAGGTGCTCGAAGACATATTACCCATGTGTATGCCAA 1688
Db      425  AGTTTACATAAAGCATTTGATGCTGTTTCTTAAACATATCTATCTTAATACATGTTAA 366
Qy      1689 GATGTTTAGCTAG 1701
Db      365  GATGTTTAGAAAG 353

RESULT 15
LOCUS   BZ797681
DEFINITION   PUGBP40TB ZM_0.6_1.0_KB Zea mays genomic clone ZMBTA331G08,
genomic survey sequence.
ACCESSION   BZ797681
VERSION     BZ797681.1 GI:29000287
KEYWORDS    GSS.
SOURCE      Zea mays
ORGANISM    Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 949)
AUTHORS    Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
TITLE      Maize Genomics Consortium
JOURNAL    Unpublished (2003)
COMMENT    Other_GSSs: PUGBP40TD
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
FEATURES   Location/Qualifiers
            source
            1..949
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /strain="B73"
            /db_xref="taxon:4577"
            /clone="ZMBTA331G08"
            /clone_lib="ZM_0.6_1.0_KB"
            /note="Vector: PCR4-TOPO; Site:1: EcoRI; 0.6-1.0 kb high
            CoT selected genomic DNA library"

ORIGIN
Query Match 6.5%; Score 175; DB 8; Length 949;
Best Local Similarity 64.7%; Pred. No. 6e-36;
Matches 319; Conservative 0; Mismatches 150; Indels 24; Gaps 3;

Qy      1232 GAGCTCATCATTTATATAAGCATGGTGATACCAAAATTAATCTCTTTTGAGCATGTTTATA 1291
```

```

Db      271  GGGCTCATAGATATCAAACTTCATGCCACTAAAGTTCTCCTTTTGACCTTATGTATG 330
Qy      1292  TTTAGGAGGACGTTTACCTGTTGAGGTAATCTGAACGCTAATAAATCGGCTAAGCAAA 1351
Db      331  GAAATGAAGCCATCTTGCCTTATAGAAATAAATCTAACCACTATTAGGTTTGCCAGGCAAA 390
Qy      1352  ATAAATTTATCACCTGTTGATTCTTAACAATTTGATGATGGACAATATTGATGAGGTGACTG 1411
Db      391  ATAAATTTAGATGTCGATGAGTACCATGATTTAATGATGGACAATATTGACGAGGTGACCG 450
Qy      1412  ACAATGATTTGAAGGCTTTAAAGGAGATTGAGAAGGATATAA-----TCT 1455
Db      451  ATAAAGGTTGATGGCTTTTGAAGGAAATTCGAAGGACAACTTATAGTCGCCAAGGTCT 510
Qy      1456  ACAATAAAATGTAAGAGAAAGCAATCAAGTGTGAGATCTGGTGTGGAAGACTATTT 1515
Db      511  ACAATAAGAAAGTTAAAGACTAAATCATTTTCAGGTTGGAGATTTGGTGTGGAAGACAATGT 570
Qy      1516  TGCTCTTTGGGGTAAAGACAAACAGTTTATAGTGGCTCAAAATTTGGGAGGCCCA 1575
Db      571  TGCC-ATTGAGATTAGGATTACAGTTTGACAAGTGGTCTCCAAGTTTGAAGGACT 629
Qy      1576  TGCAAGATTGTTAAAGTAATTTGTTTGGAT-----TGACGGAGGCAITTCAGGTGAT 1628
Db      630  TATAAGGTTGTACATGTAATCCCTGGTAATGCTTATATGTTGGAACATTACAGGAAT 689
Qy      1629  CATCTACCTAGAGCTCTCAATGGAGGTGCTCGAAGACATATTACCCTATGTGTATGGCAA 1688
Db      690  AGTTTACATAAAGCATTTGAATGGTTCGTTTCCCTAAACCAATACTATCCTAATACATGGTAA 749
Qy      1689  GATGTTTAGCTAG 1701
Db      750  GATGTTTAGAAG 762

```

Search completed: May 9, 2005, 04:46:22
 Job time : 8601 secs